20874.0 96e1s001

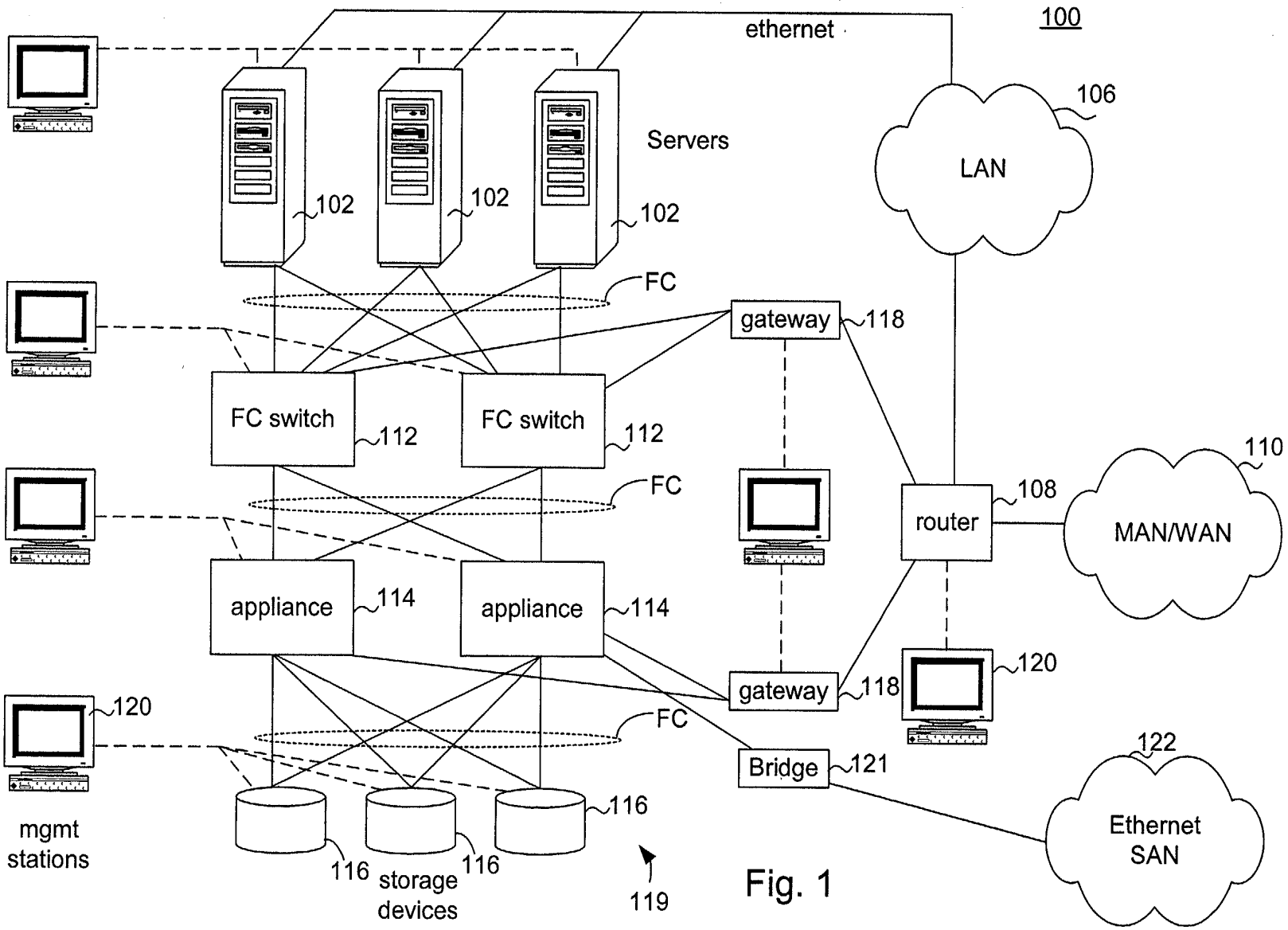


Fig. 1

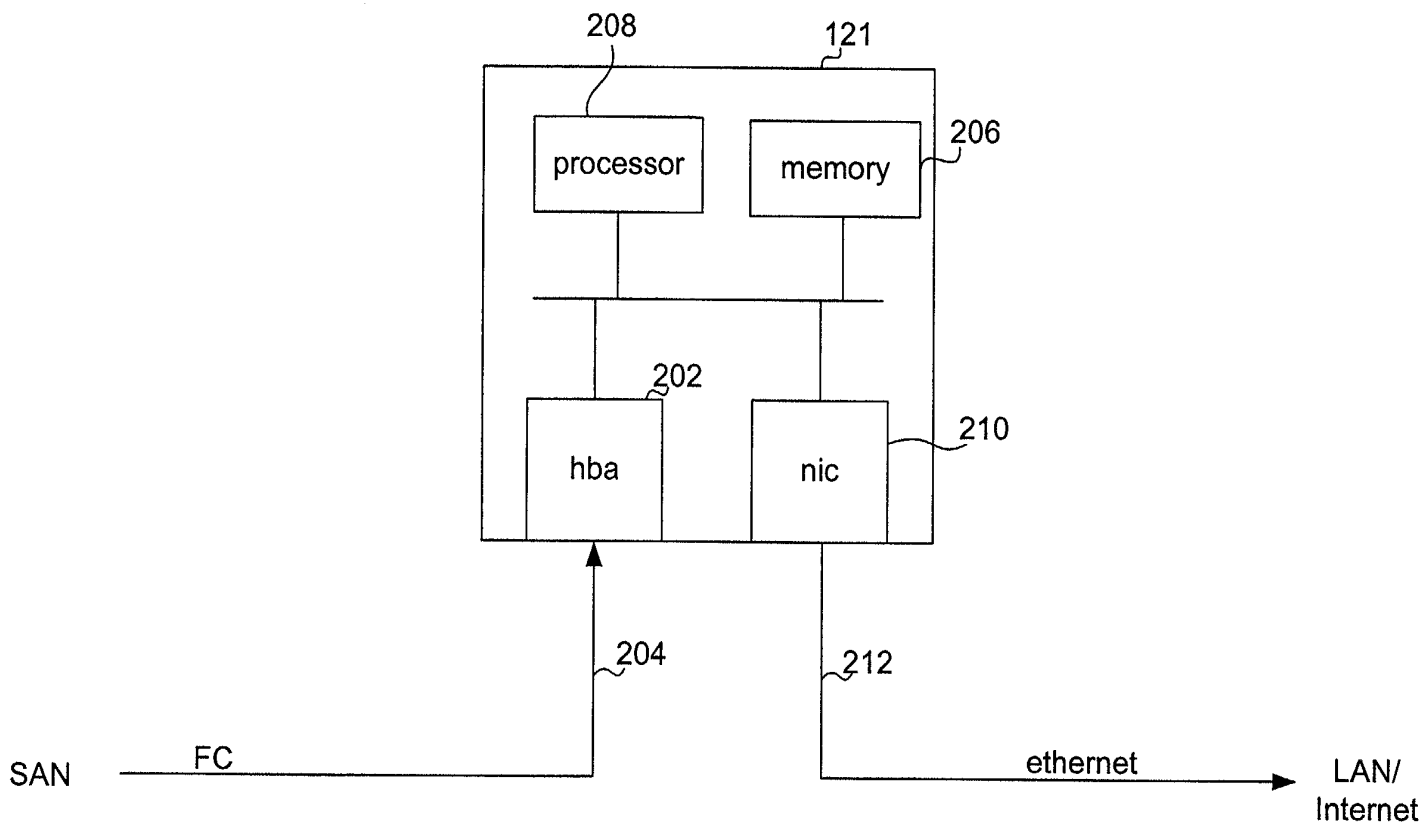

208
121
processor
memory
206
202
hba
nic
210
204
212
SAN
FC
ethernet
LAN/
Internet

Fig. 2

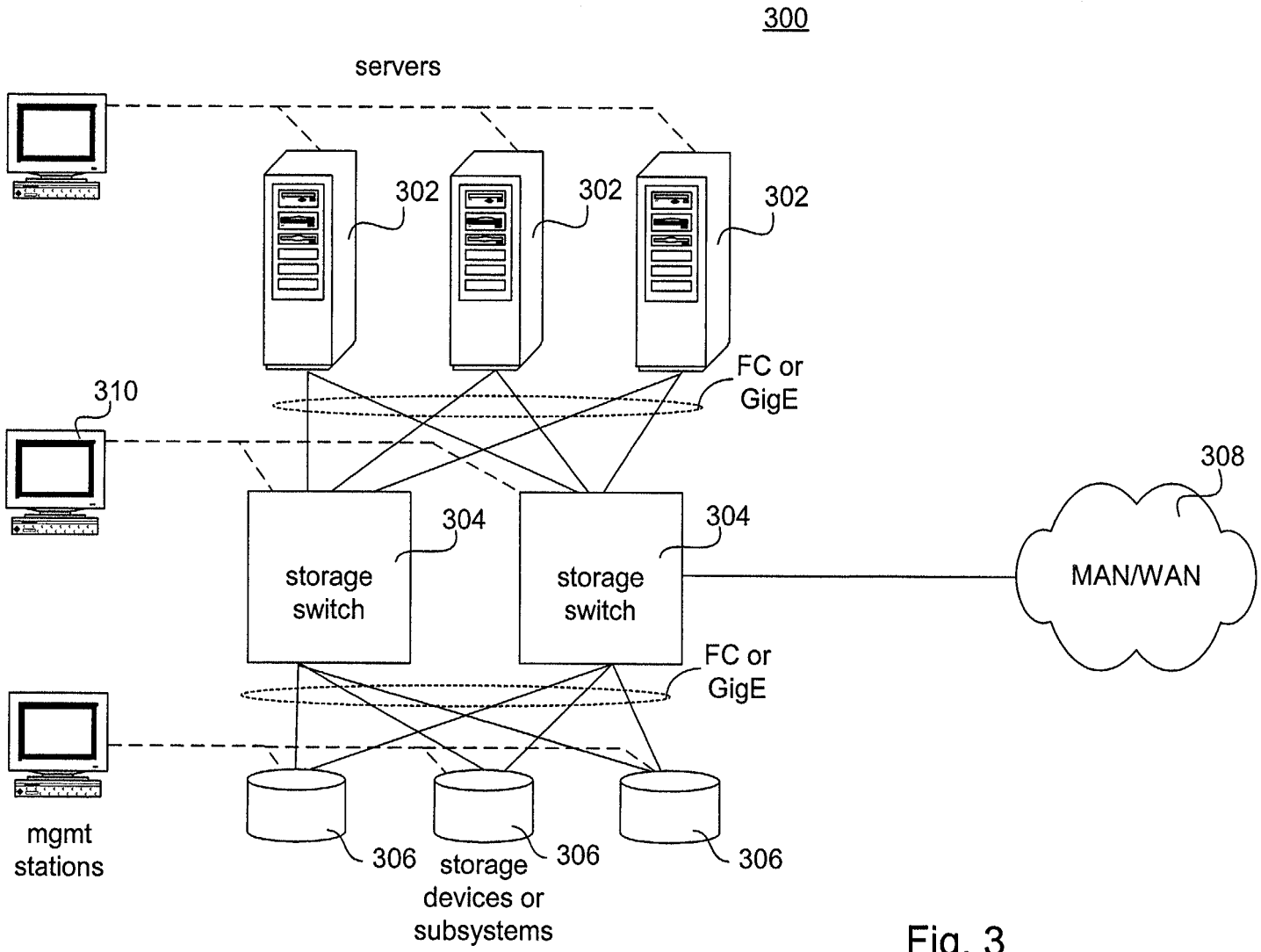

300
servers
302
302
302
FC or
GigE
310
304
304
storage
switch
storage
switch
308
MAN/WAN
FC or
GigE
mgmt
stations
306
306
306
storage
devices or
subsystems

Fig. 3

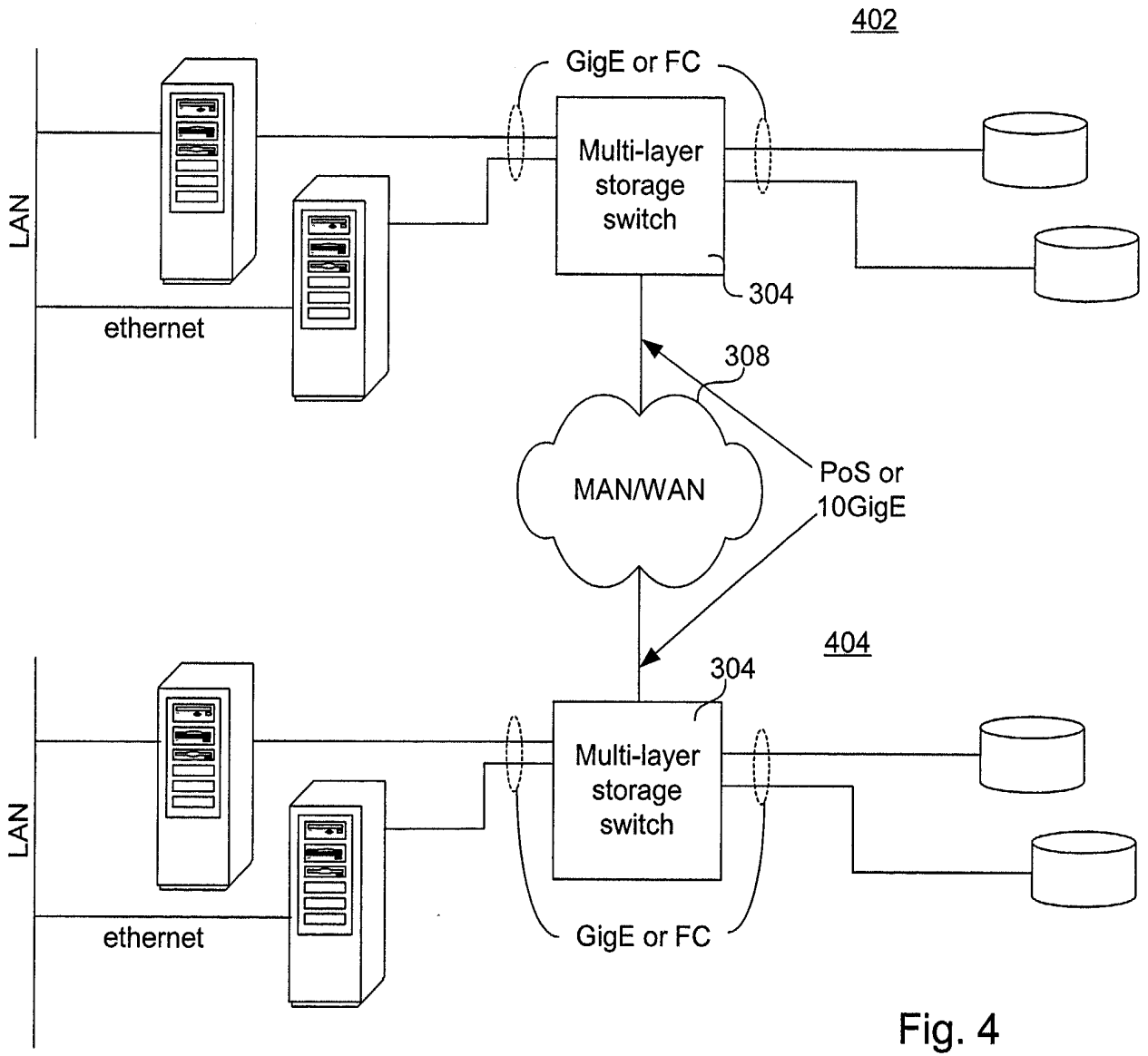

402
GigE or FC
Multi-layer storage switch
304
LAN
ethernet
308
MAN/WAN
PoS or 10GigE
404
304
Multi-layer storage switch
LAN
ethernet
GigE or FC

Fig. 4

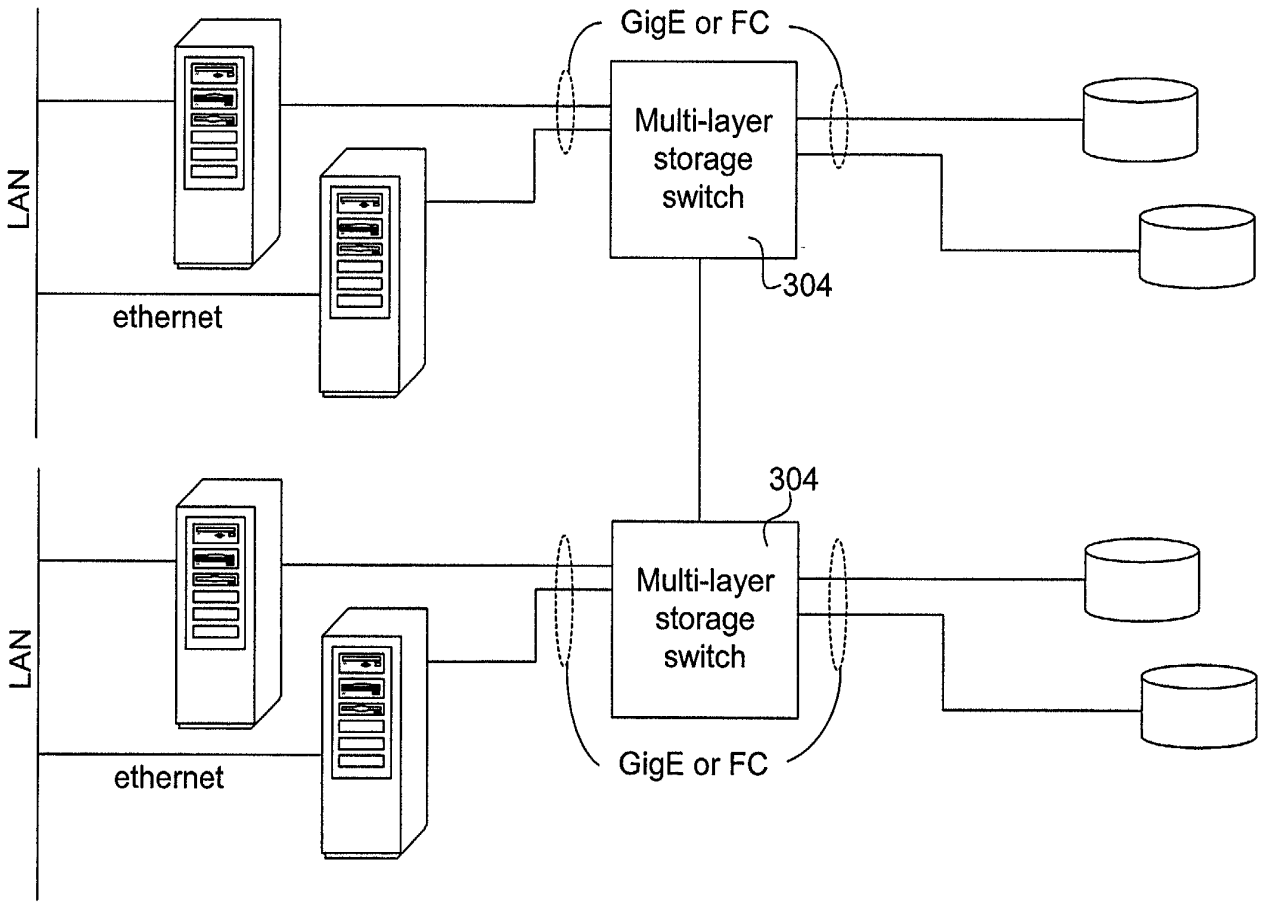

GigE or FC
Multi-layer storage switch
304
LAN
ethernet
304
Multi-layer storage switch
LAN
ethernet
GigE or FC

Fig. 5

Page 6 of 38
Applicant: Santosh C. Lolayekar et al.
Title: VIRTUALIZATION IN A STORAGE SYSTEM
Attorney Docket No.: MARA-01005US0 SBS
Attorney: Sarah Barone Schwartz Express Mail No.: EL670728891US

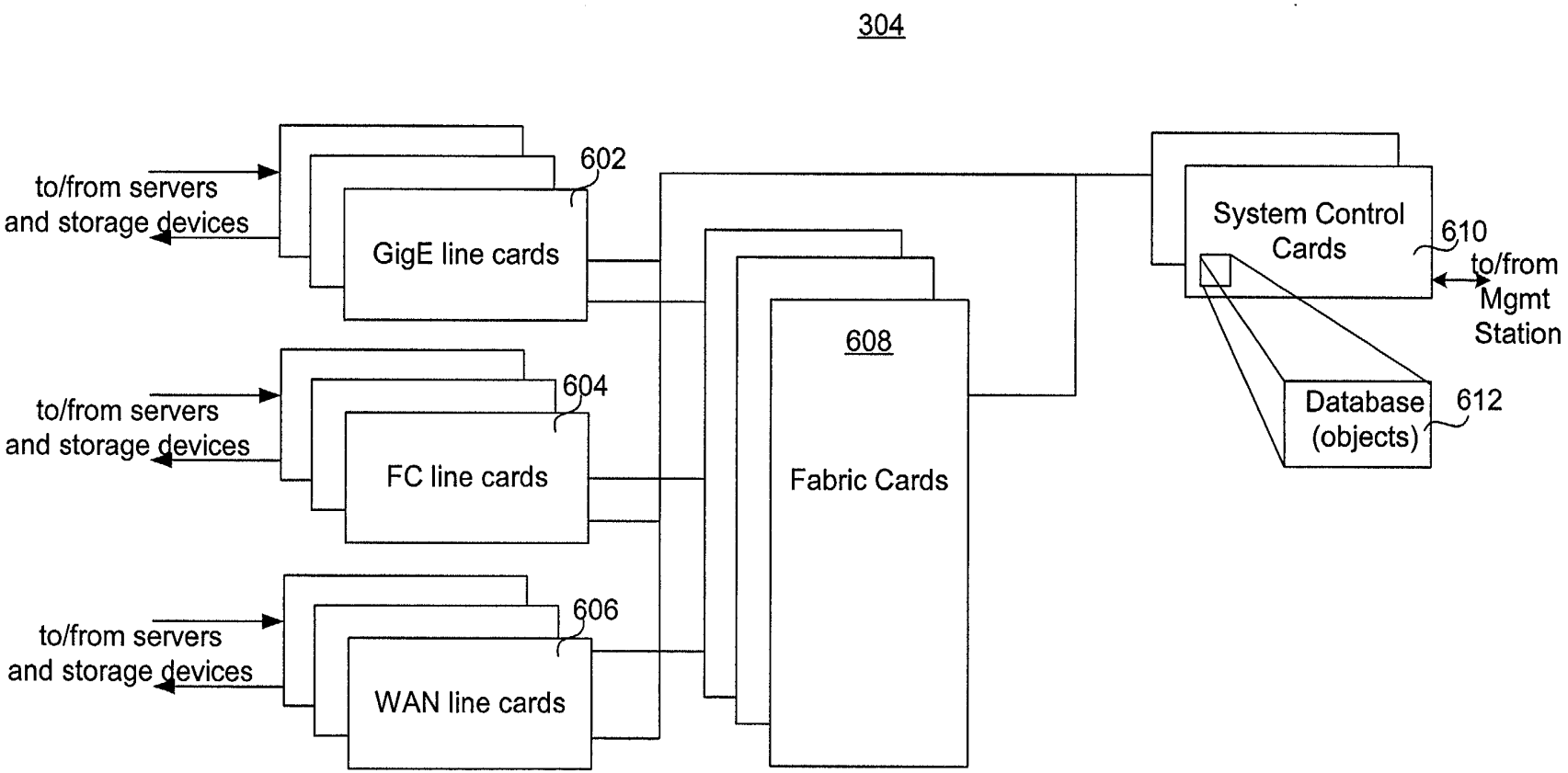


Fig. 6

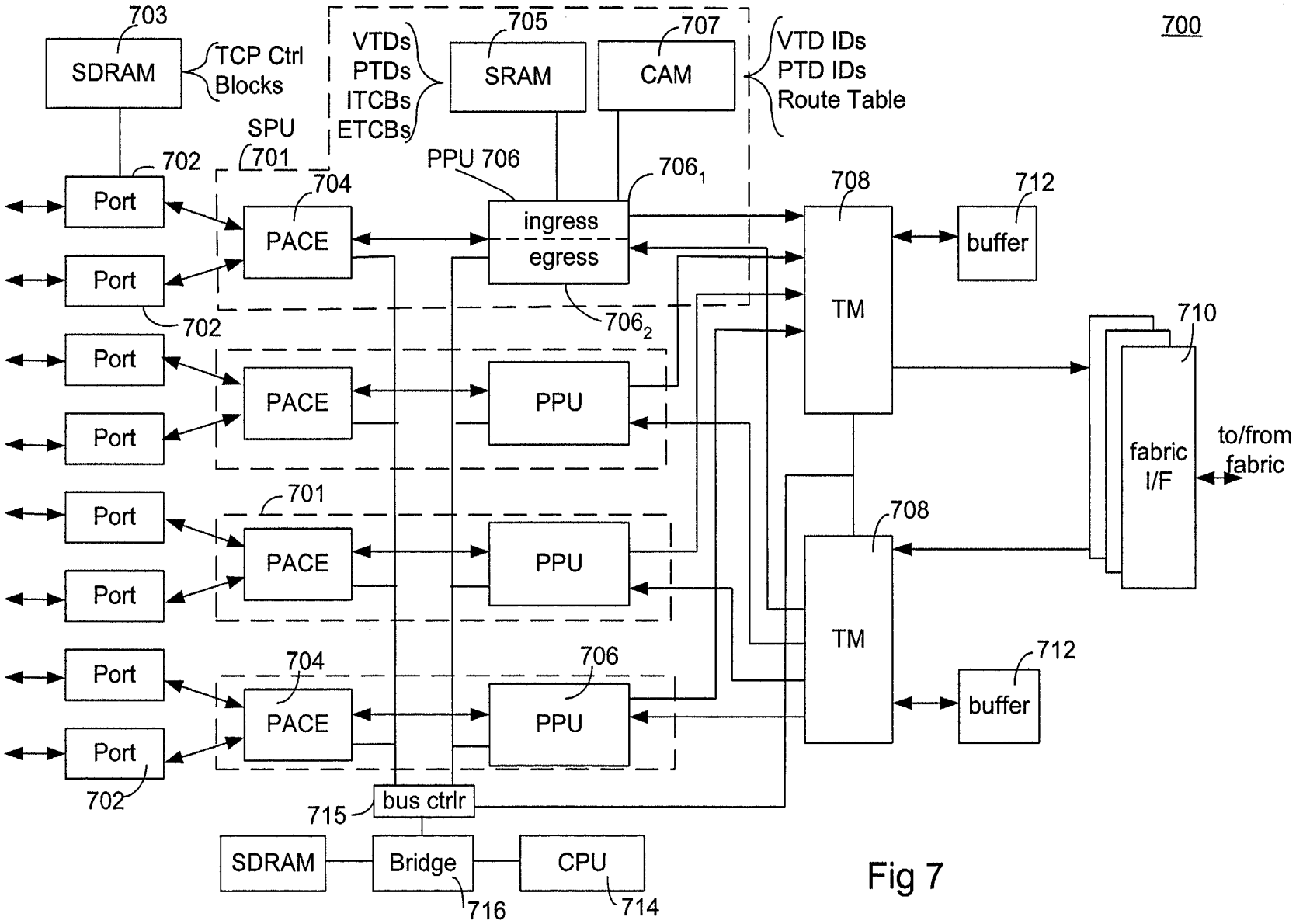

700
703
SDRAM
TCP Ctrl Blocks
VTDs
PTDs
ITCBs
ETCBs
705
SRAM
707
CAM
VTD IDs
PTD IDs
Route Table
SPU
701
PPU 706
706_1
706_2
ingress
egress
702
Port
704
PACE
PPU
708
TM
712
buffer
710
fabric I/F
to/from fabric
bus ctrlr
715
SDRAM
Bridge
716
CPU
714

Fig 7

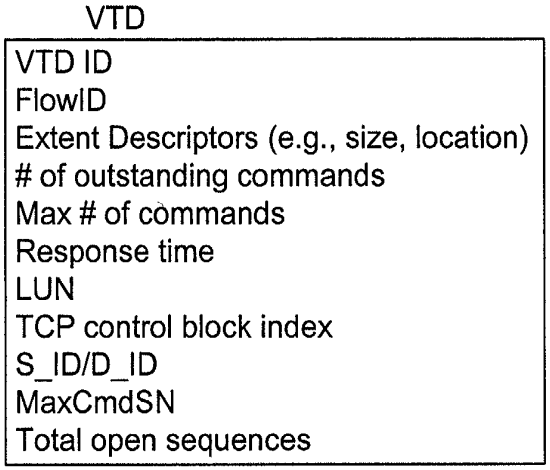
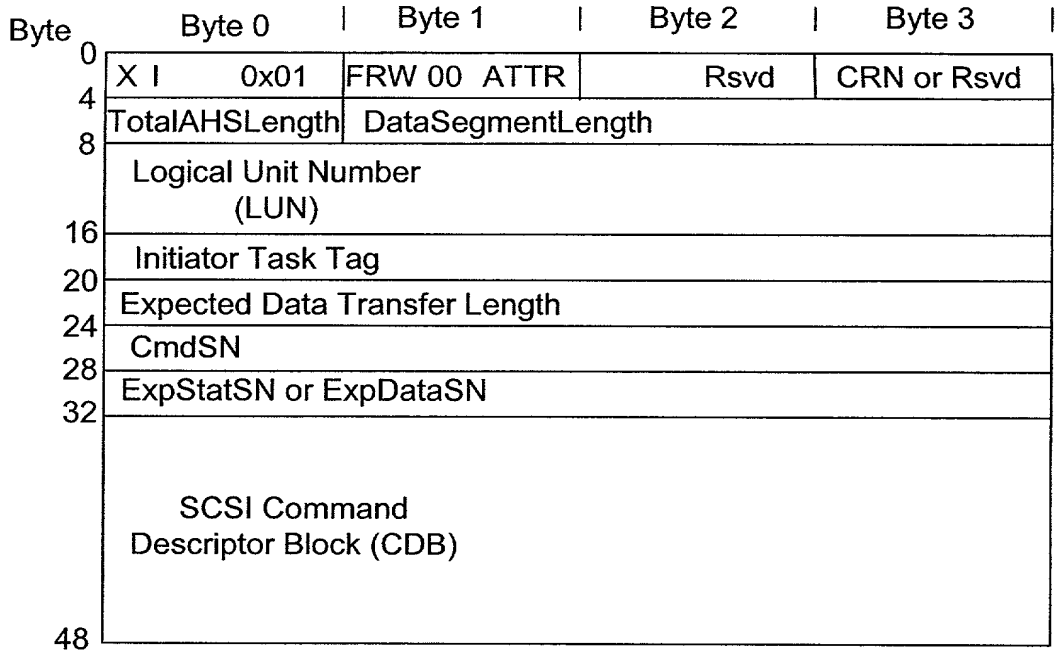
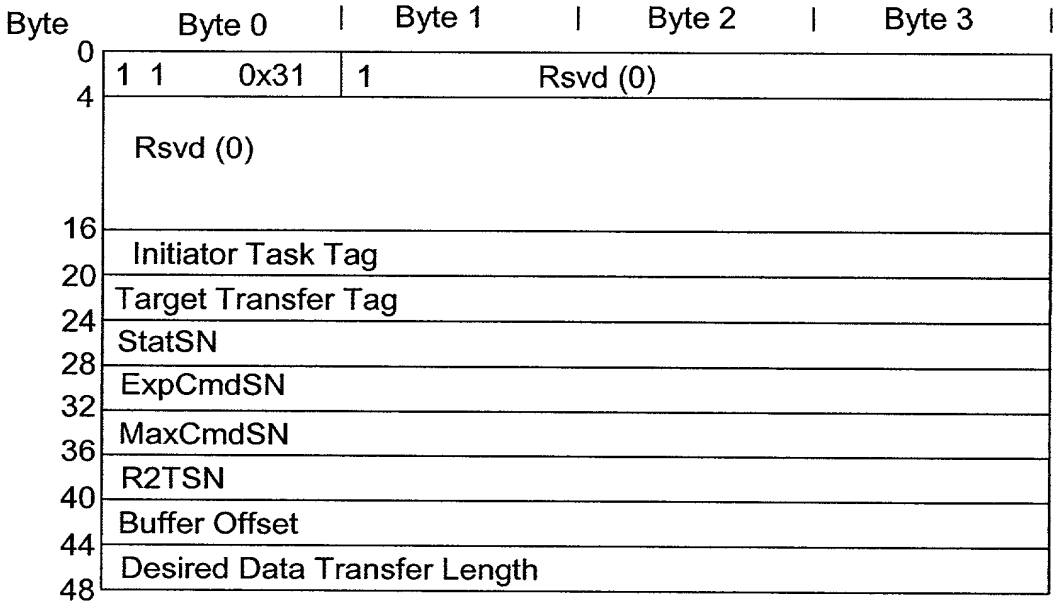

VTD
VTD ID
FlowID
Extent Descriptors (e.g., size, location)
# of outstanding commands
Max # of commands
Response time
LUN
TCP control block index
S_ID/D_ID
MaxCmdSN
Total open sequences

Fig. 7a



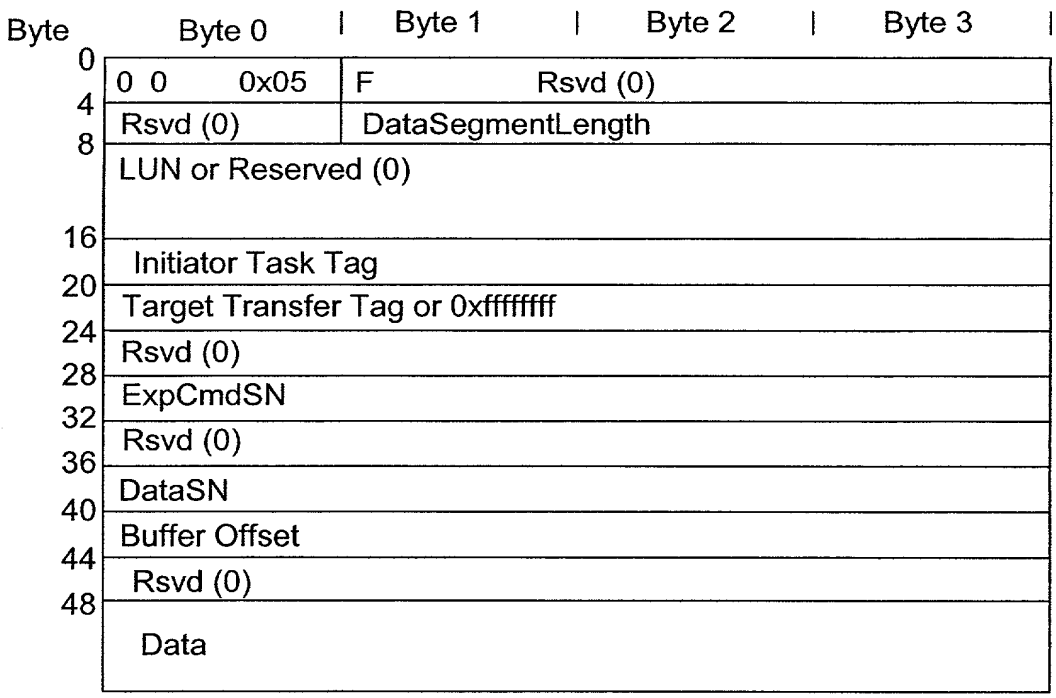
iSCSI Command PDU

Fig. 8a



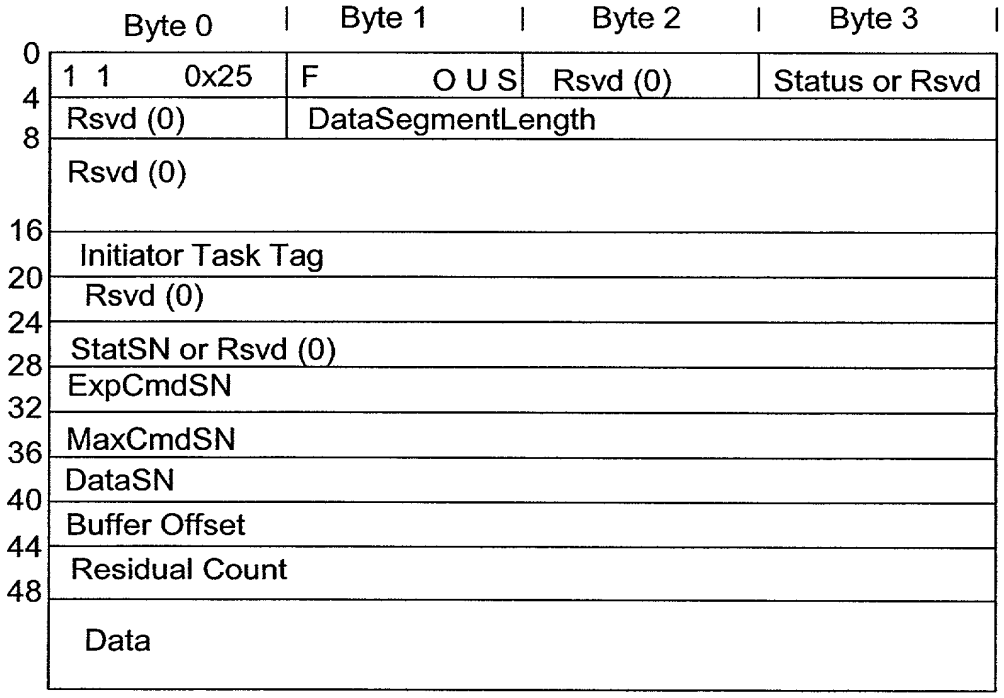
iSCSI R2T PDU

Fig. 8b



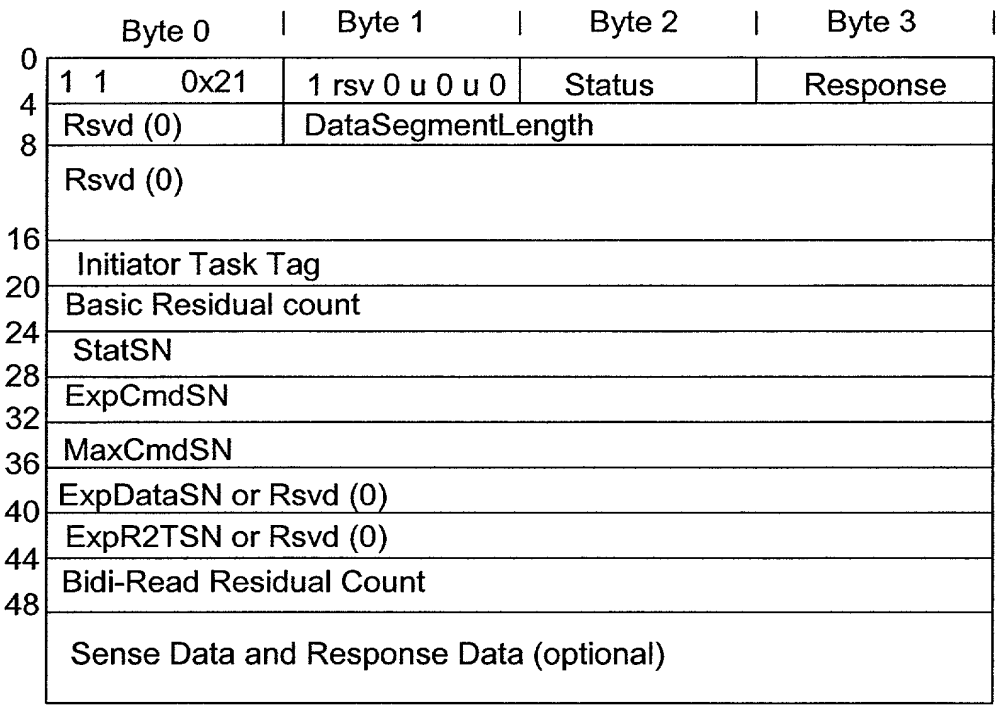
iSCSI Write Data PDU

Fig. 8c



iSCSI Read Data PDU

Fig. 8d



iSCSI Response PDU

Fig. 8e

Bits Word	31-24	23-16	15-08	07-00
0	R_CTL	D_ID		
1	rsvd	S_ID		
2	TYPE	F_CTL		
3	SEQ_ID	DF_CTL	SEQ_CNT	
4	OX_ID		RX_ID	
5	RLTV_OFF			

FC Frame Header

Fig. 8f

Field Name	Description	Size
FCP_LUN	logical unit number	8 bytes
FCP_CNTL	control field	4 bytes
FCP_CDB	SCSI command descriptor block	16 bytes
FCP_DL	Data Length	4 bytes

FCP_CMND Payload

Fig. 8g

Field Name	Description	Size
DATA_RO	Relative offset of first byte of FCP_DATA IU that follows	4 bytes
BURST_LEN	length of FCP_DATA IU that follows	4 bytes
rsvd		4 bytes

FCP_XFR_RDY Payload

Fig. 8h

Field Name	Description	Size
rsvd		4 bytes
rsvd		4 bytes
FCP_STATUS	field validity and SCSI status	4 bytes
FCP_RESID	residual count	4 bytes
FCP_SNS_LEN	Length of FCP_SNS_INFO field	4 bytes
FCP_RSP_LEN	Length of FCP_RSP_INFO field	4 bytes
FCP_RSP_INFO	FCP response info	m bytes
FCP_SNS_INFO	FCP sense info	n bytes

FCP_RSP Payload

Fig. 8i

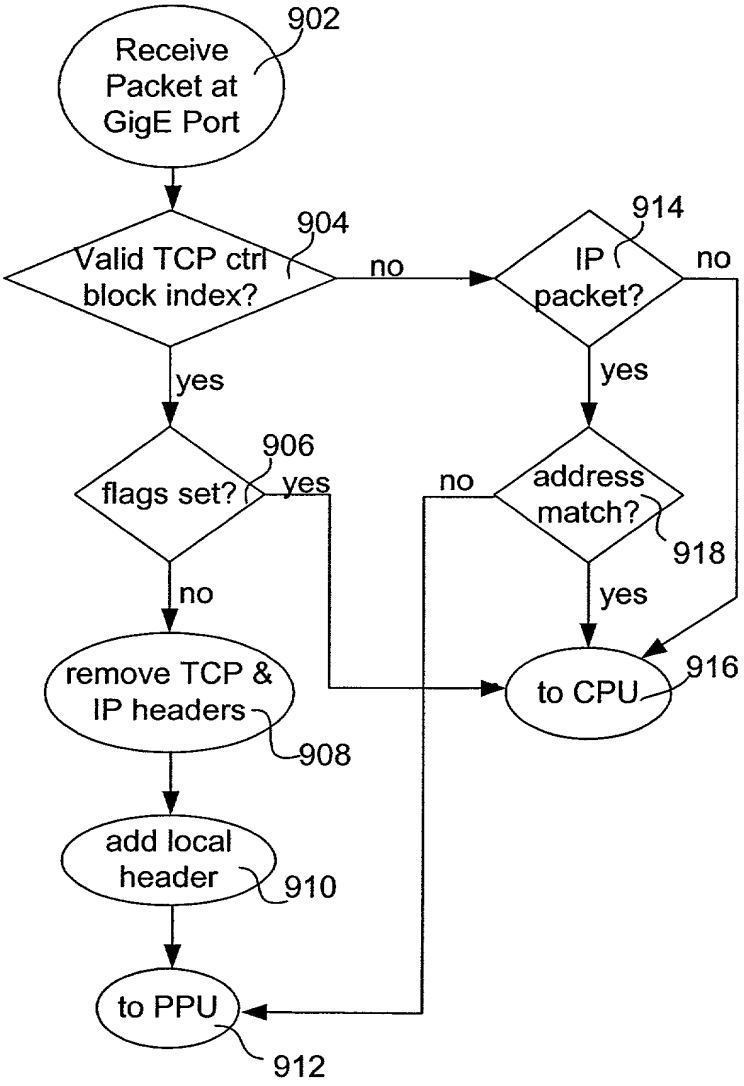


Fig 9a
(Classification - PACE - iSCSI - ingress)

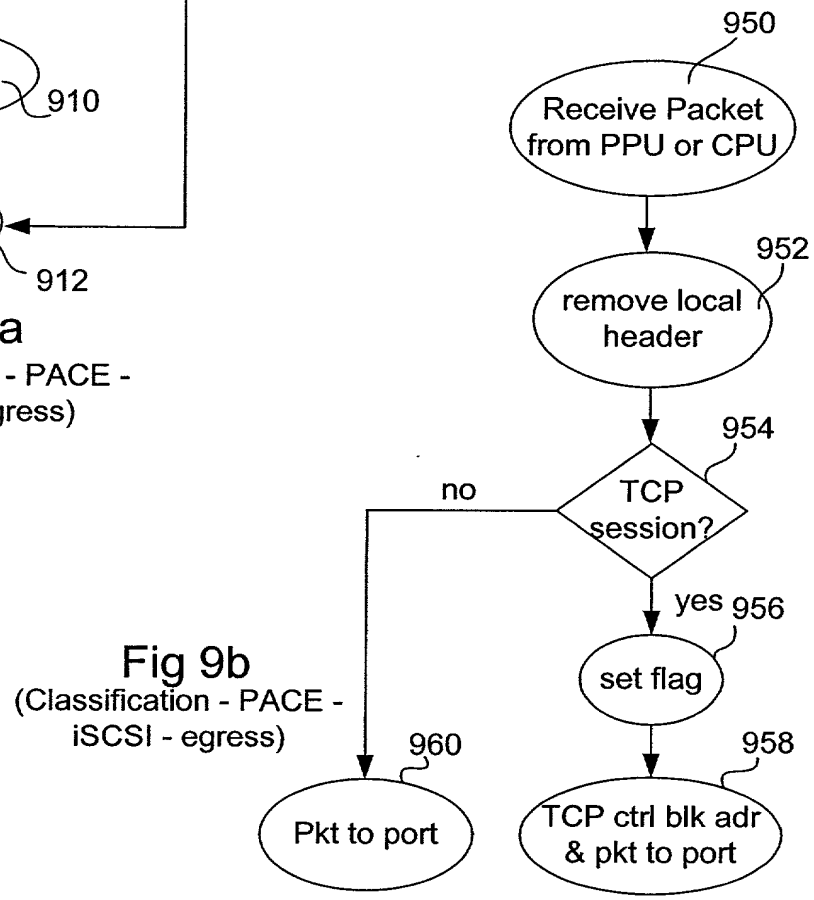


Fig 9b
(Classification - PACE - iSCSI - egress)

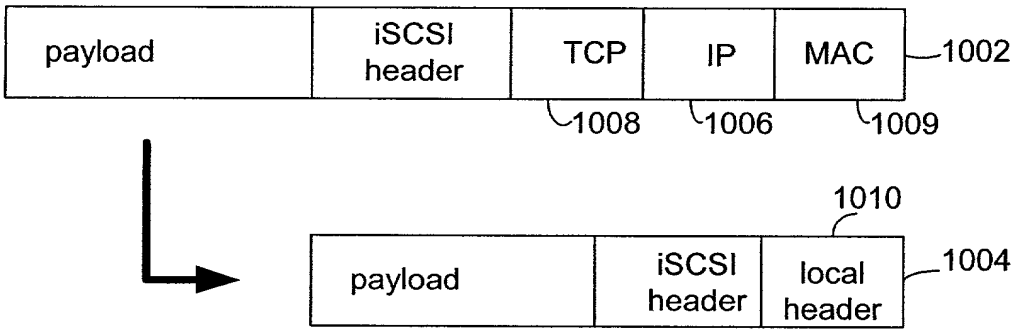


Fig. 10a

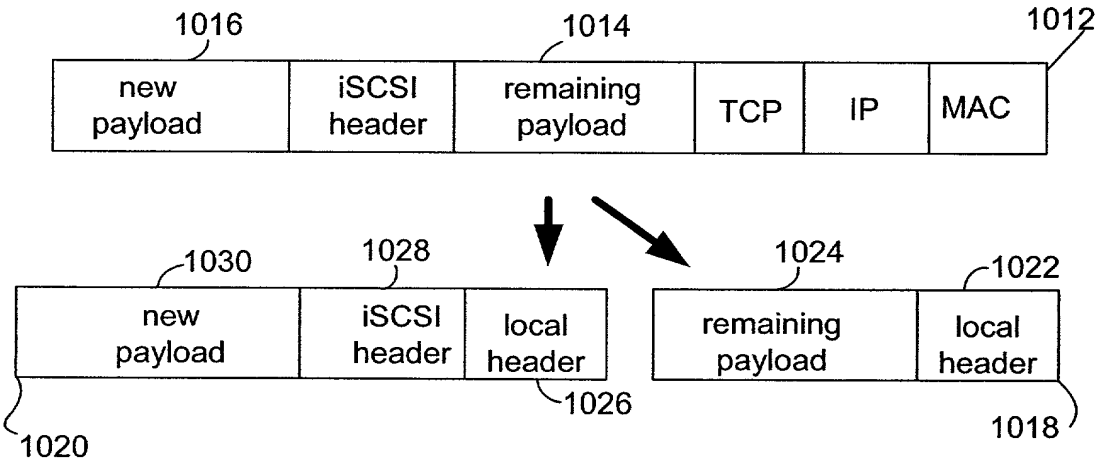


Fig. 10b

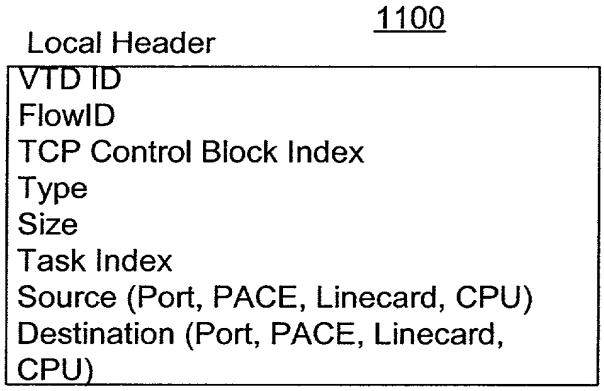

1100
Local Header
VTD ID
FlowID
TCP Control Block Index
Type
Size
Task Index
Source (Port, PACE, Linecard, CPU)
Destination (Port, PACE, Linecard, CPU)

Fig. 11

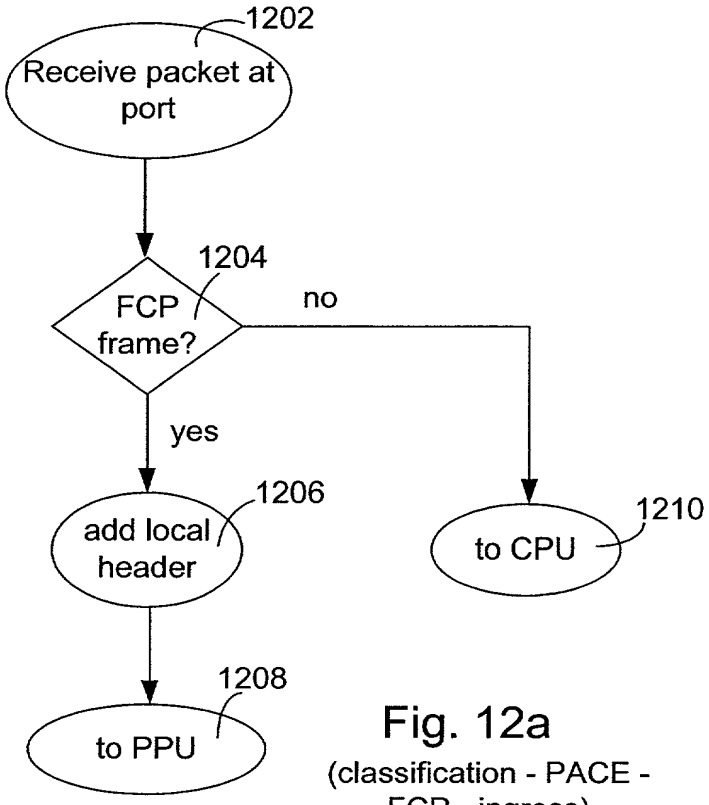


Fig. 12a
(classification - PACE - FCP - ingress)

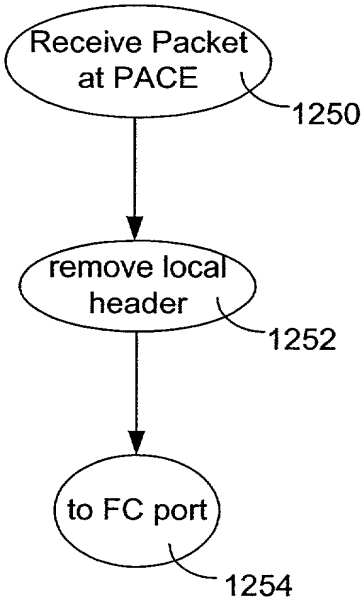


Fig. 12b
(classification - PACE - FCP - egress)

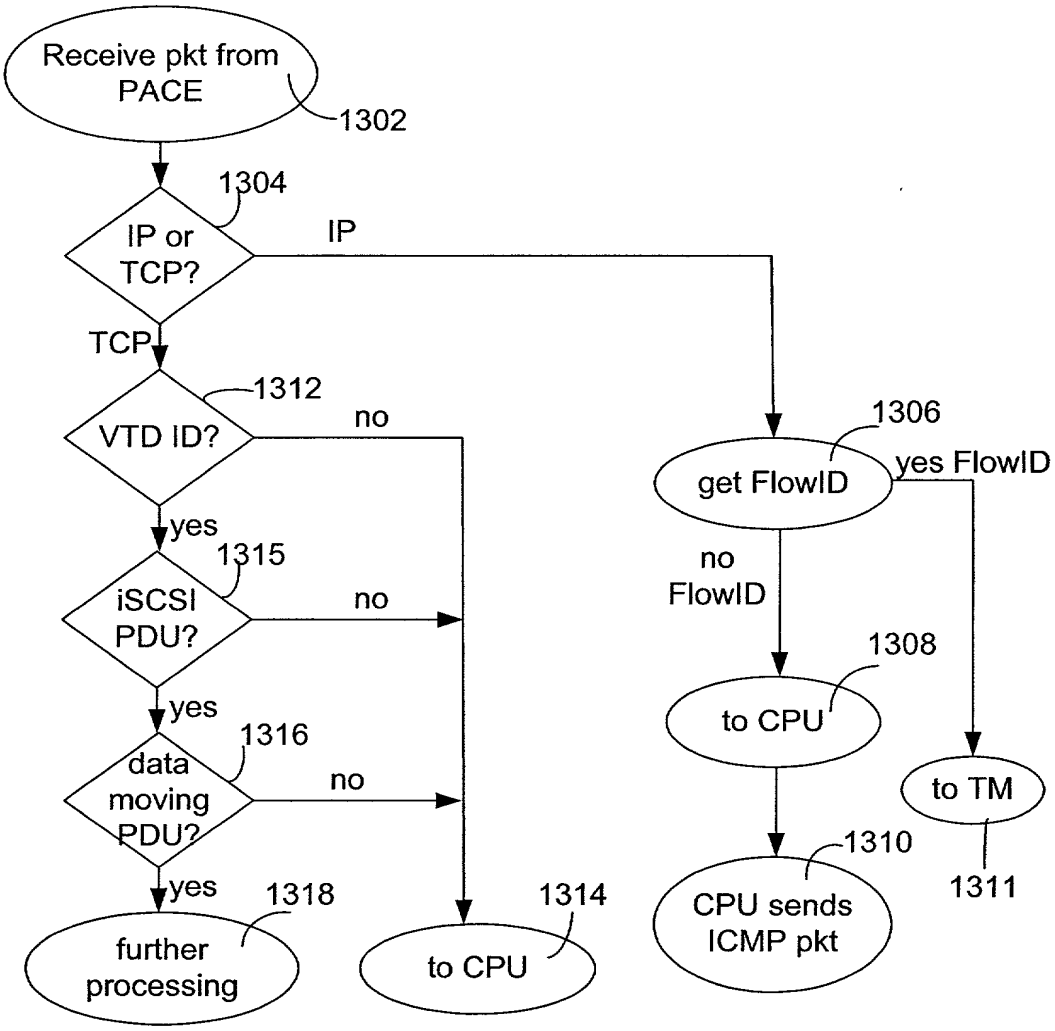


Fig. 13a
(Classification - PPU - ingress)

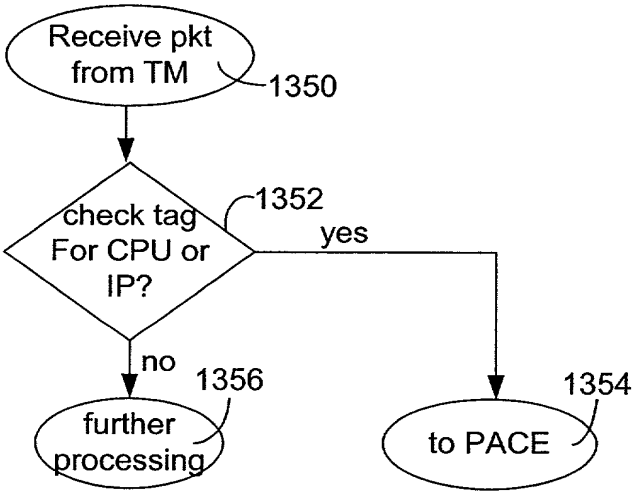


Fig. 13b
(Classification - PPU - egress)

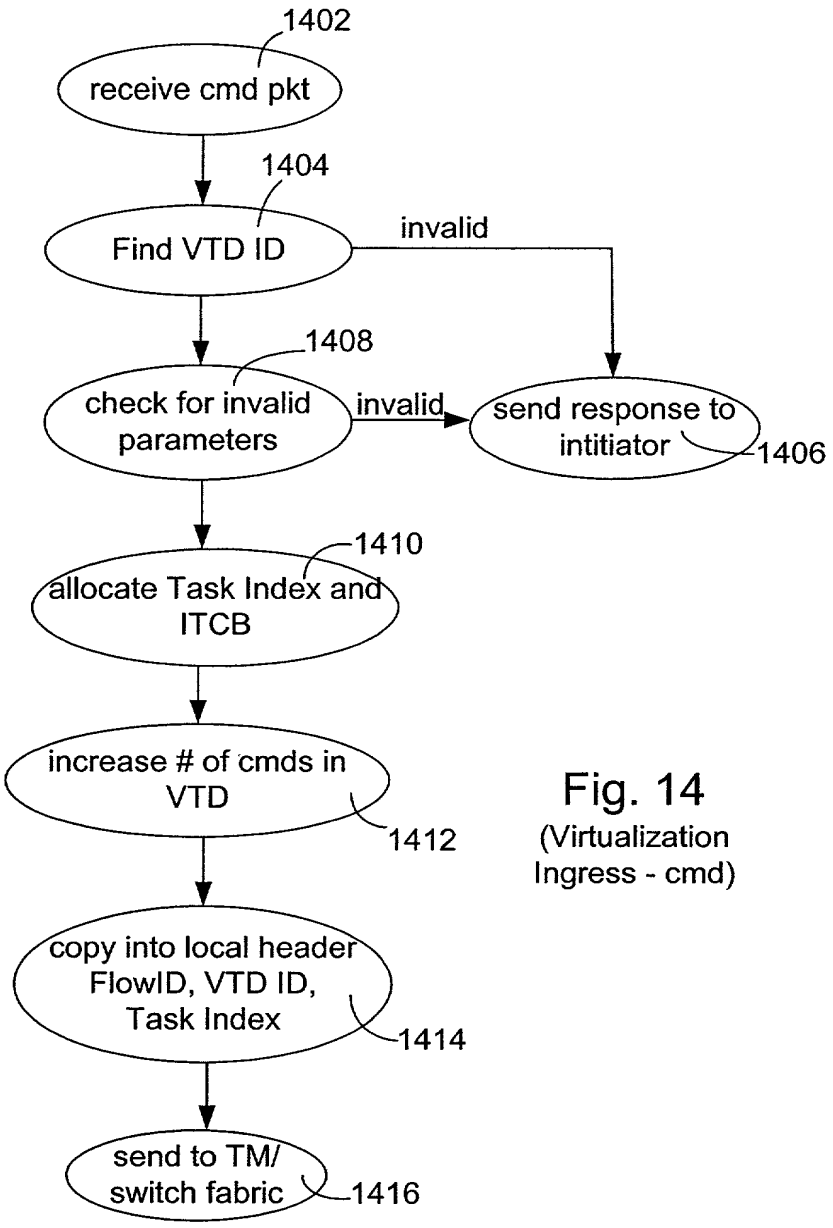


Fig. 14
(Virtualization Ingress - cmd)

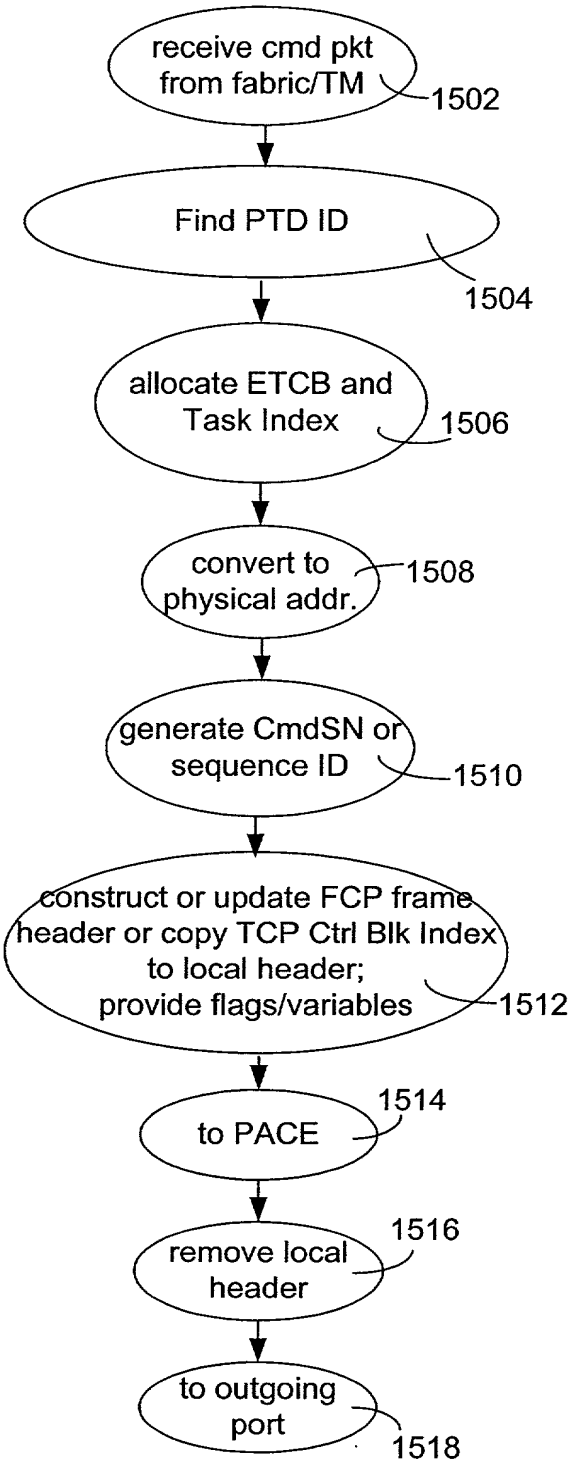


Fig. 15
(Virtualization - Egress - cmd)

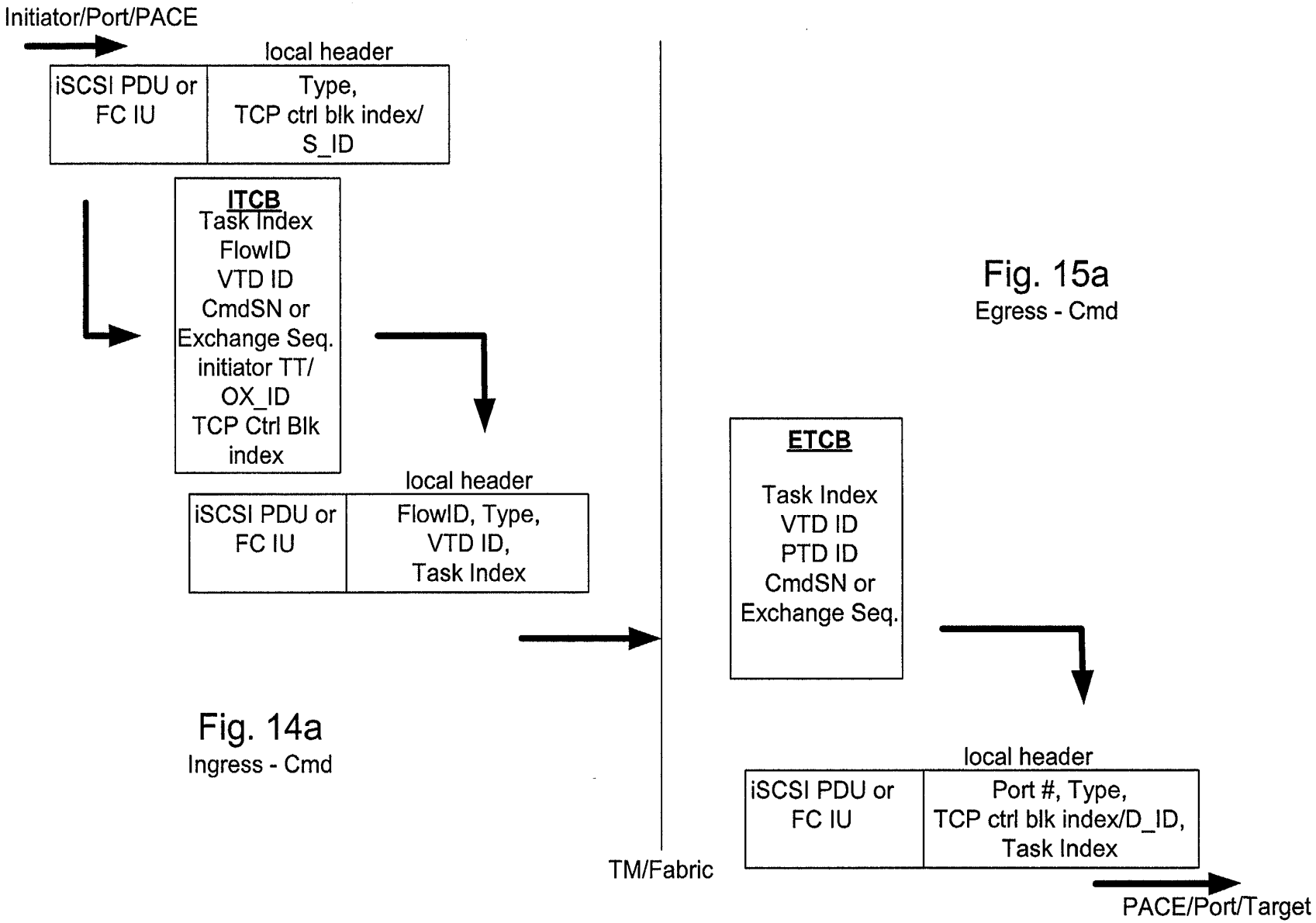


Fig. 14a
Ingress - Cmd

Fig. 15a
Egress - Cmd

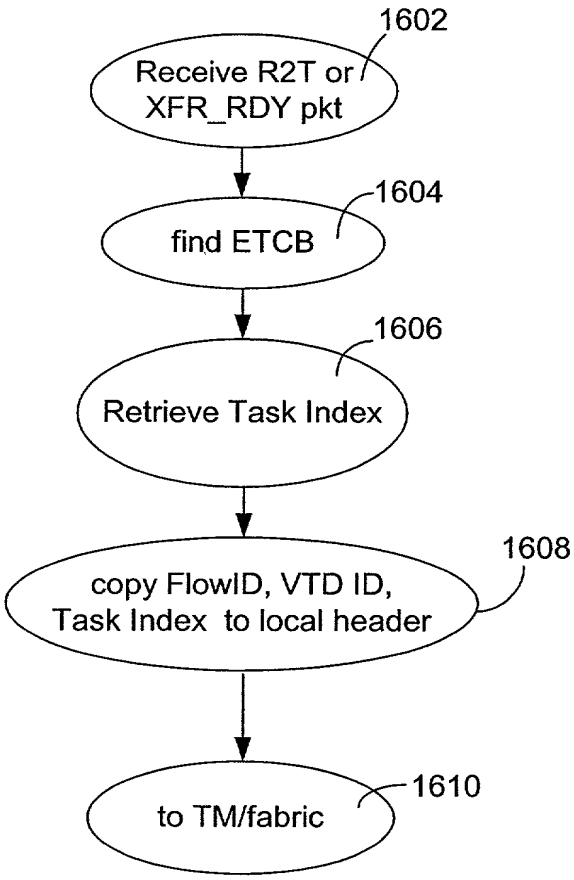


Fig. 16
(Virtualization - Ingress - R2T/XFR_RDY)

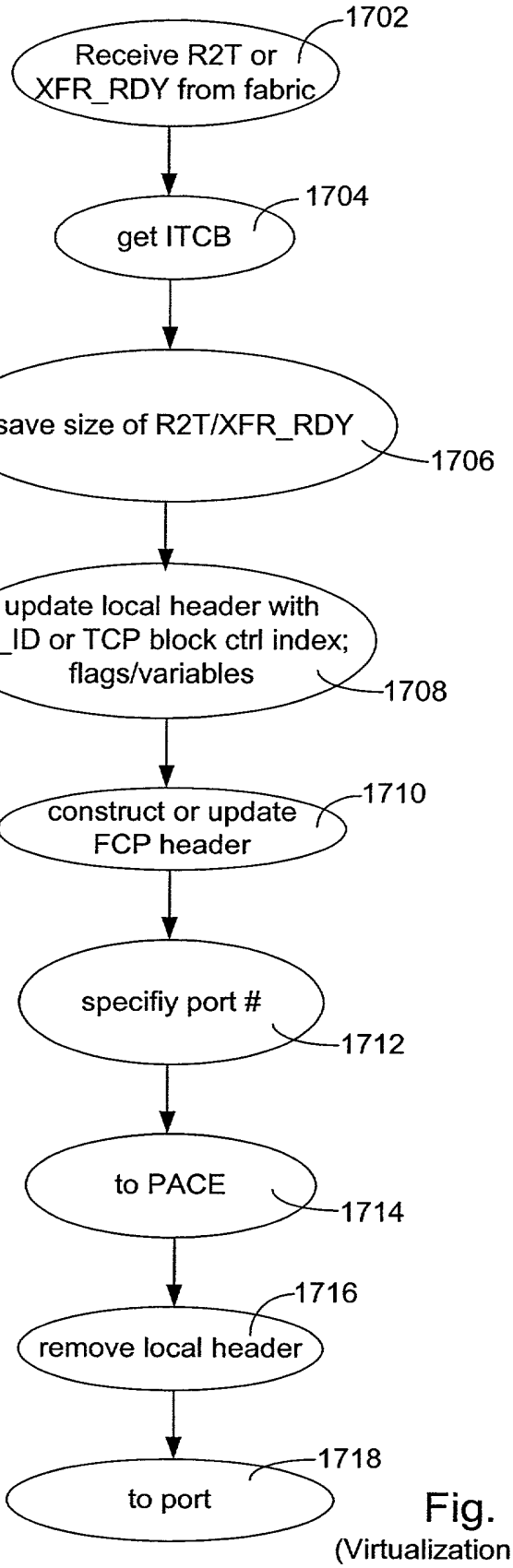


Fig. 17
(Virtualization - Egress - R2T/XFR_RDY)

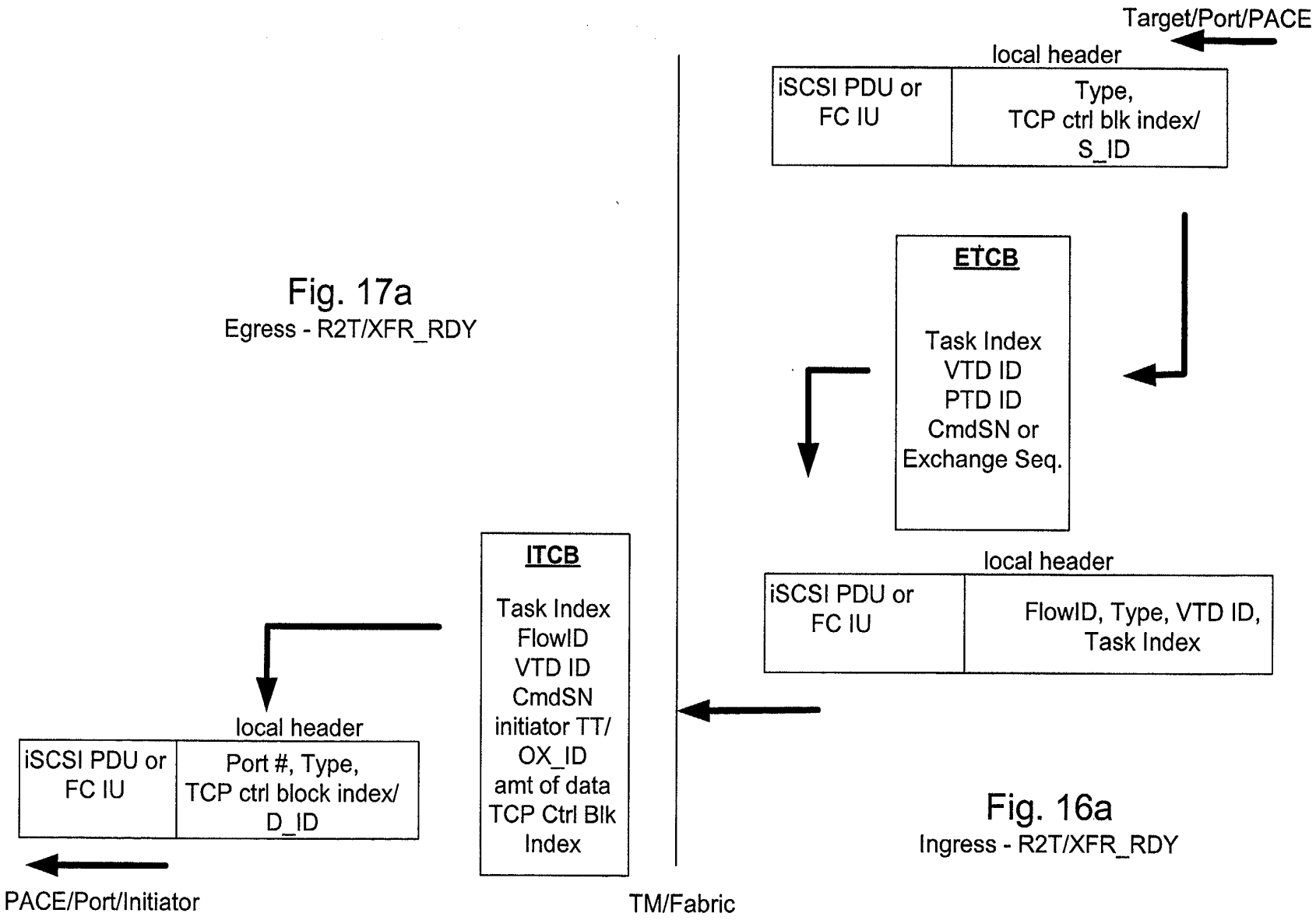


Fig. 17a
Egress - R2T/XFR_RDY

Fig. 16a
Ingress - R2T/XFR_RDY

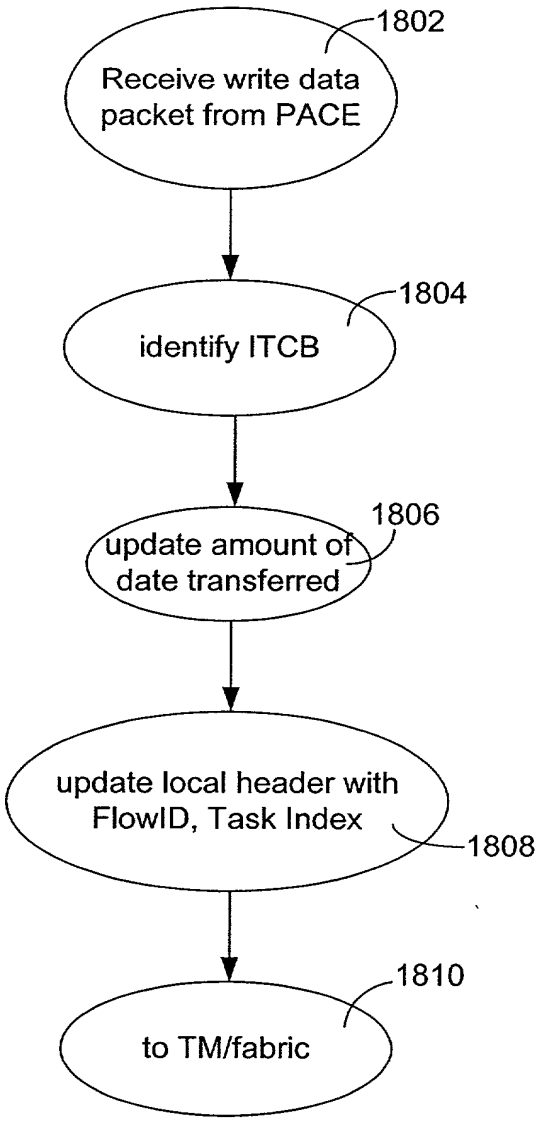


Fig. 18
(Virtualization - Ingress - write data packet)

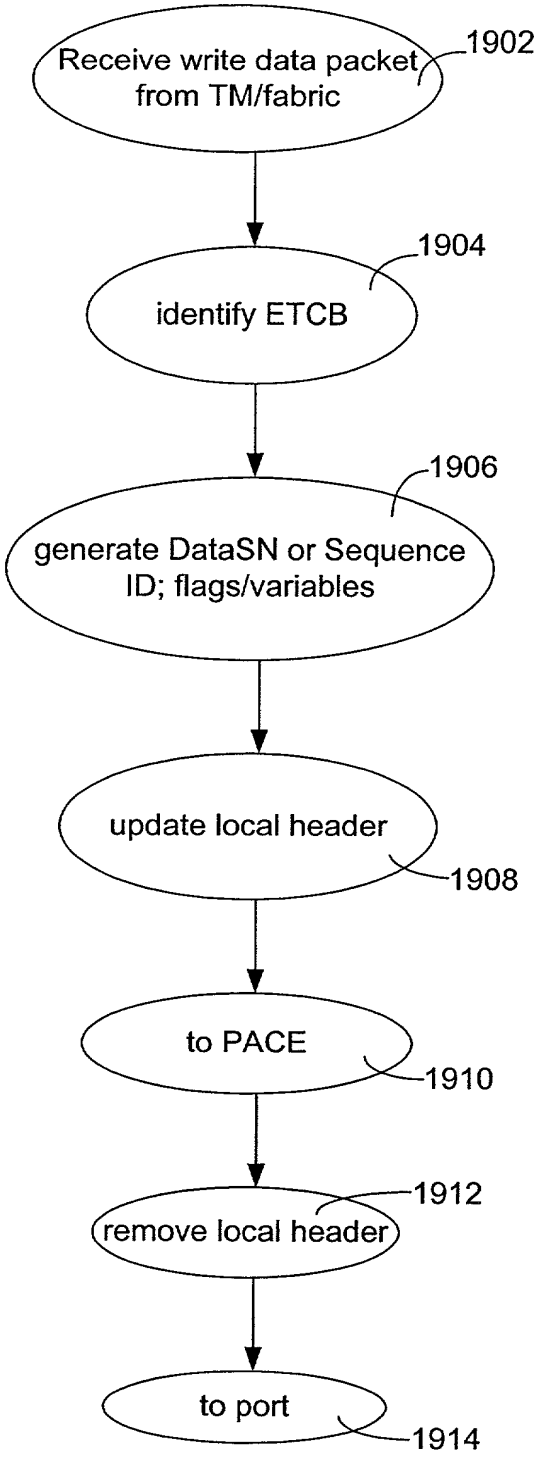


Fig. 19
(Virtualization - Egress - write data pkt)

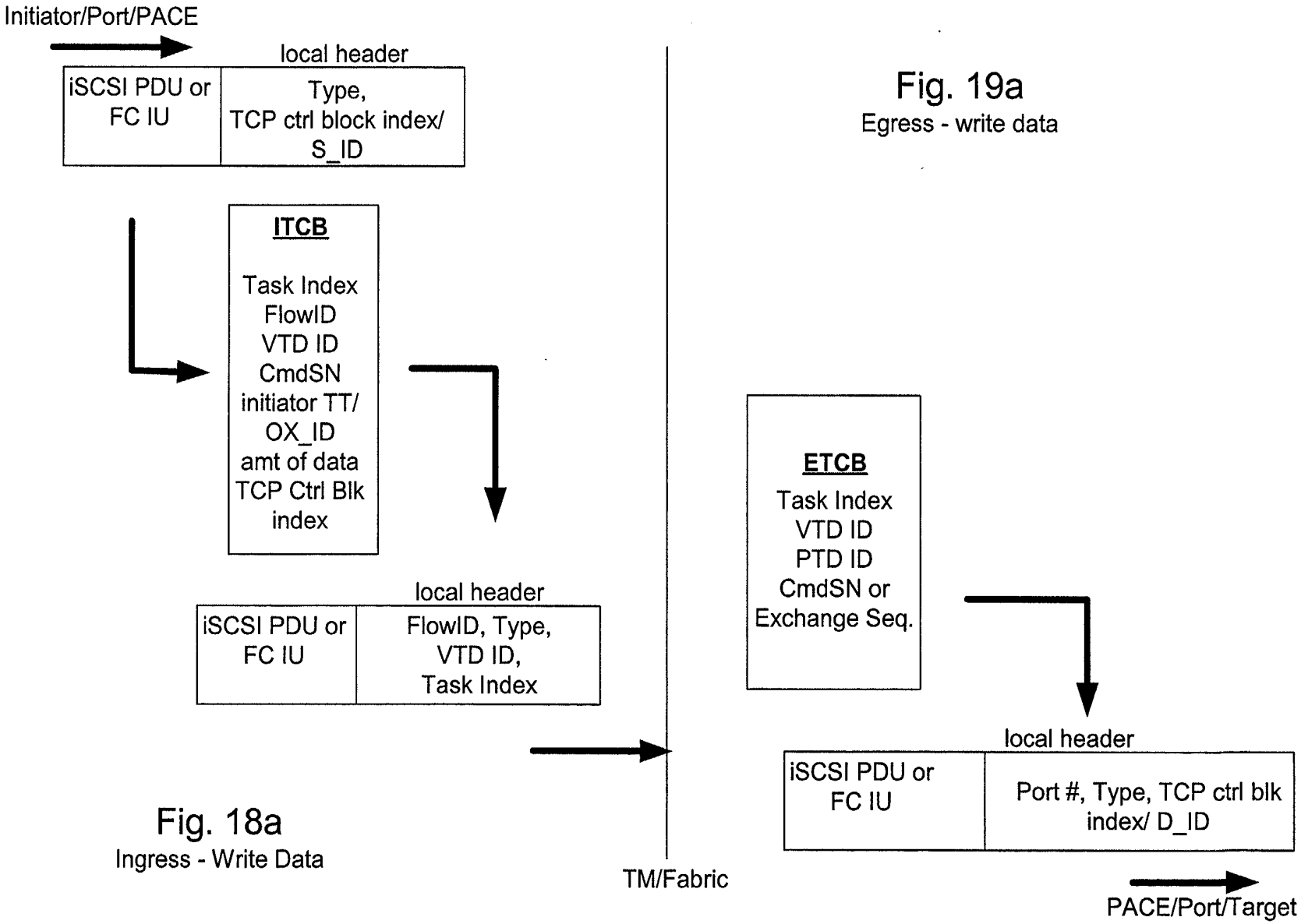


Fig. 18a
Ingress - Write Data

Fig. 19a
Egress - write data

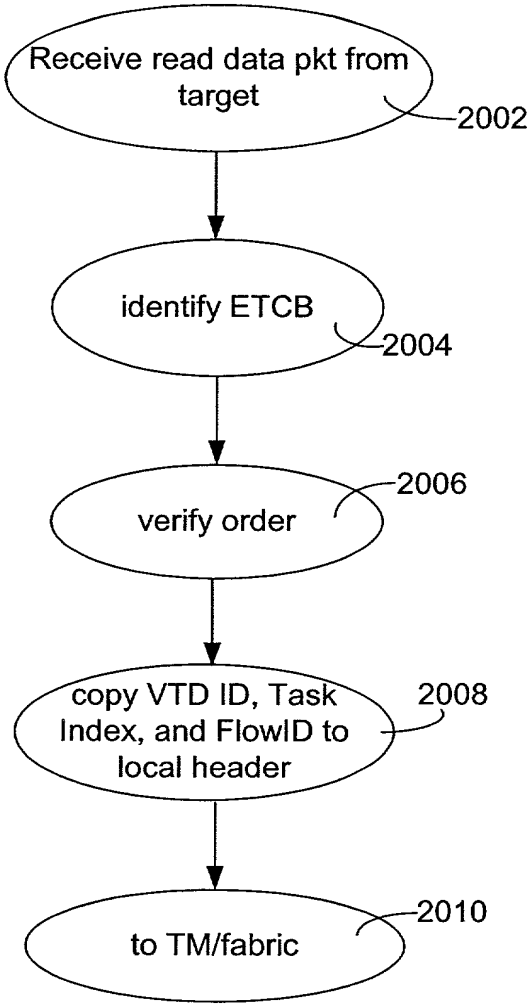


Fig. 20
(Virtualization - Ingress - Read Data pkt)

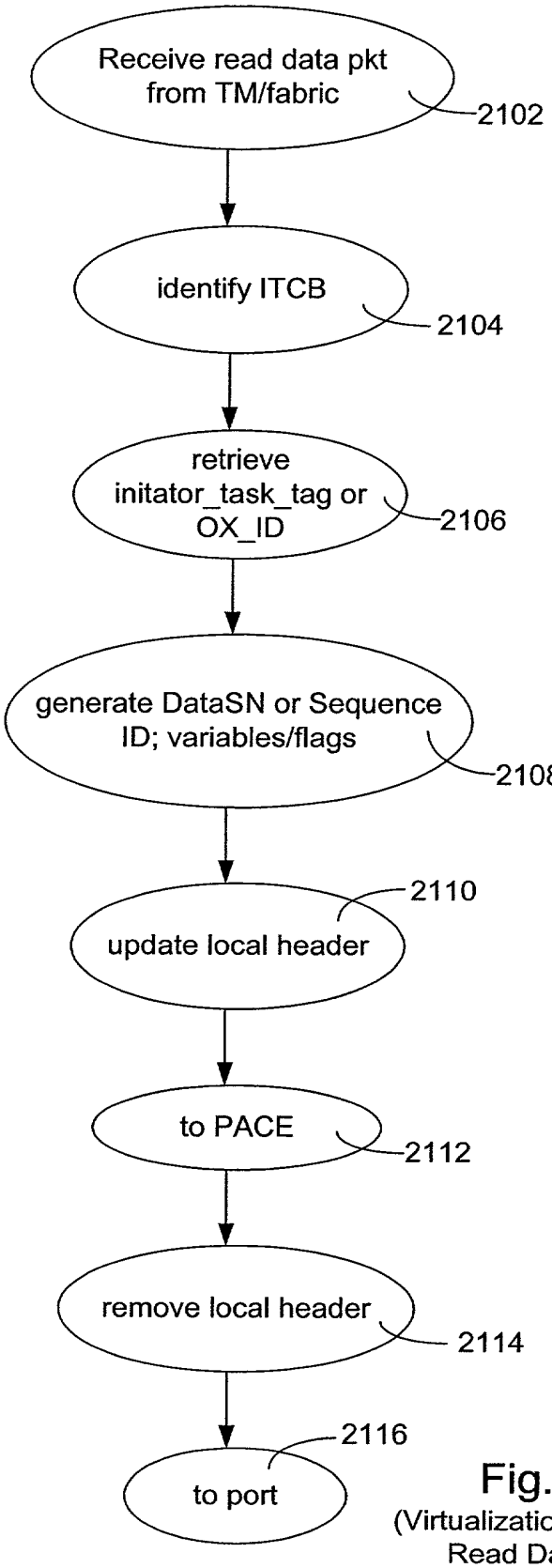


Fig. 21
(Virtualization - Egress-Read Data pkt)

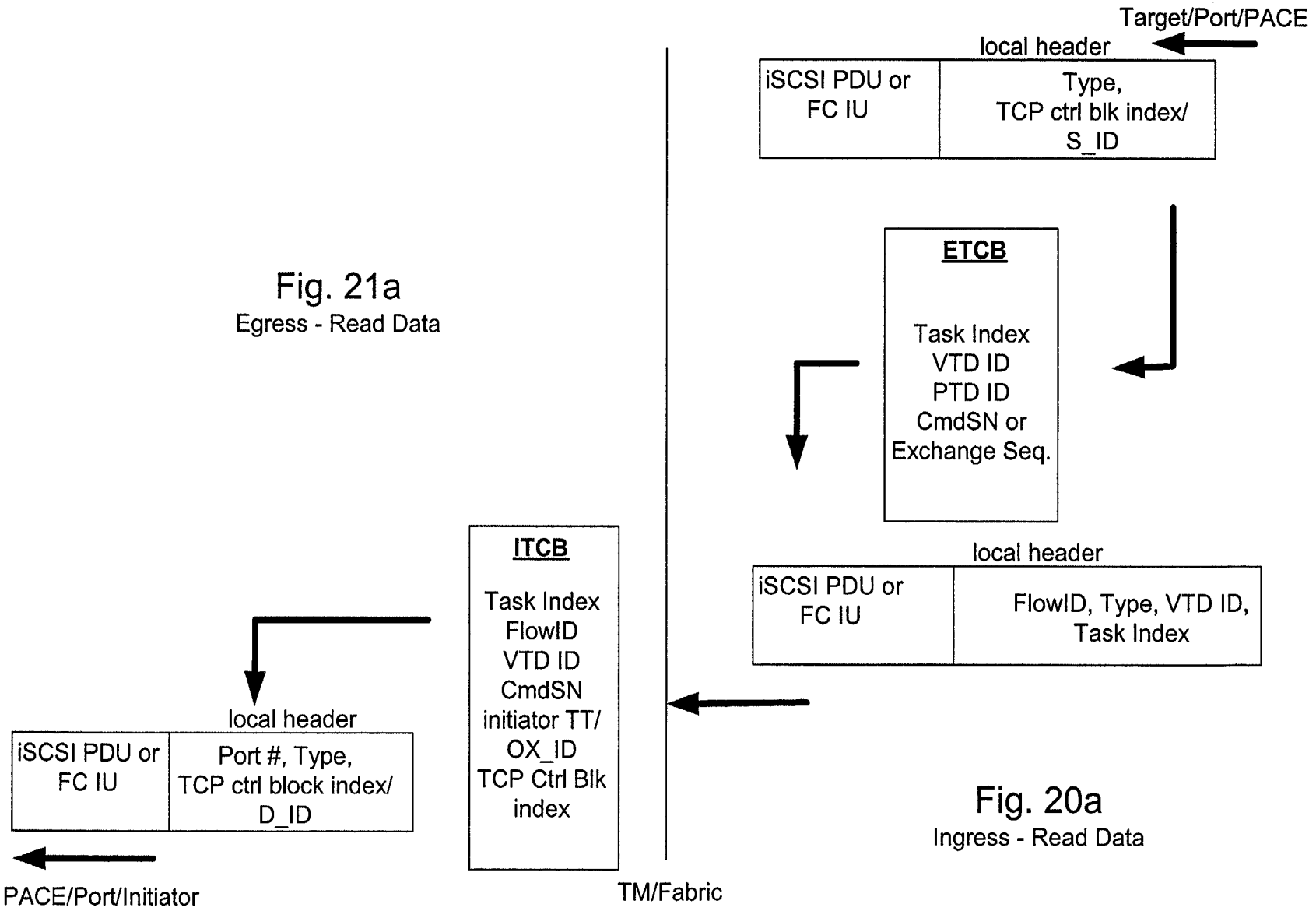


Fig. 21a
Egress - Read Data

Fig. 20a
Ingress - Read Data

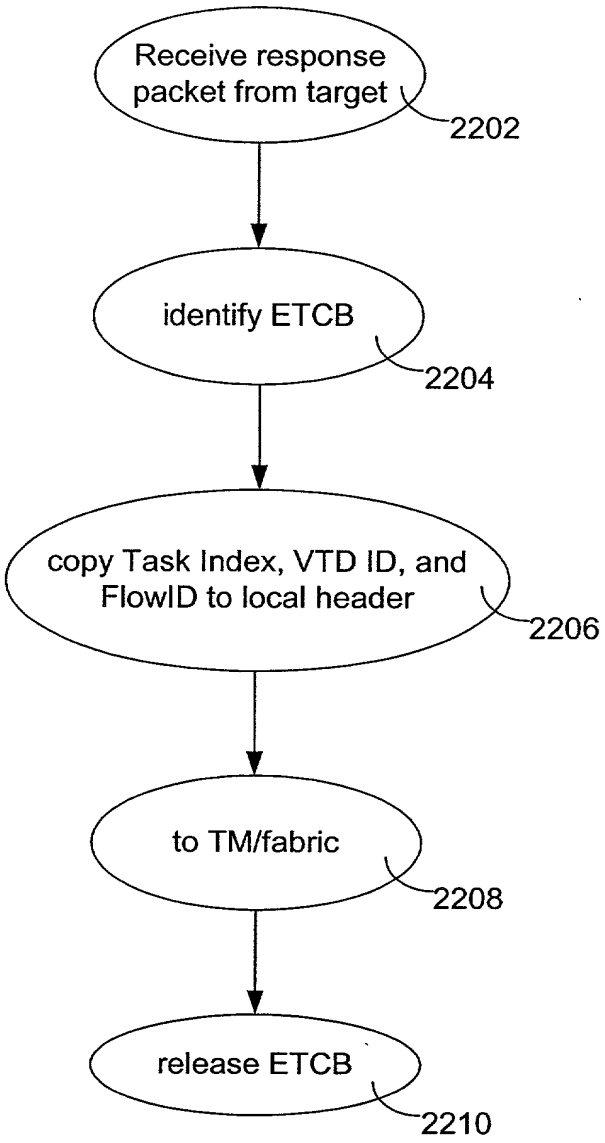


Fig. 22
(Virtualization - Ingress - response pkt)

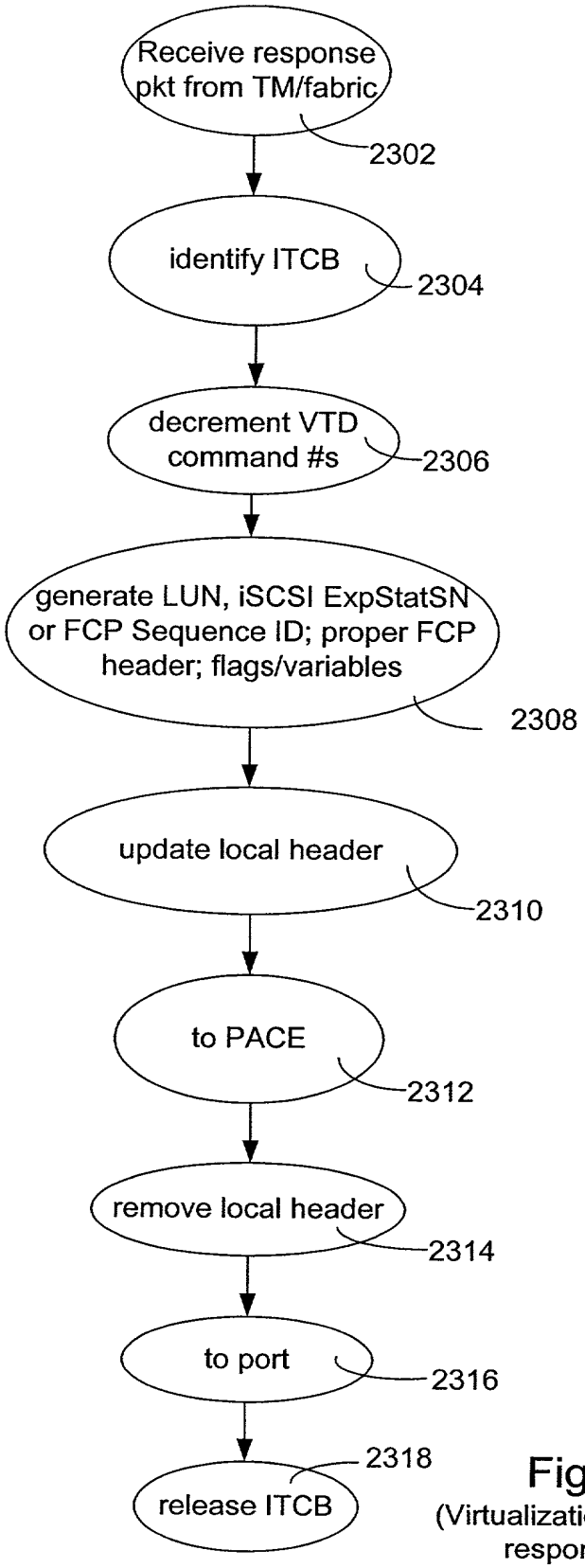


Fig. 23
(Virtualization - Egress - response pkt)

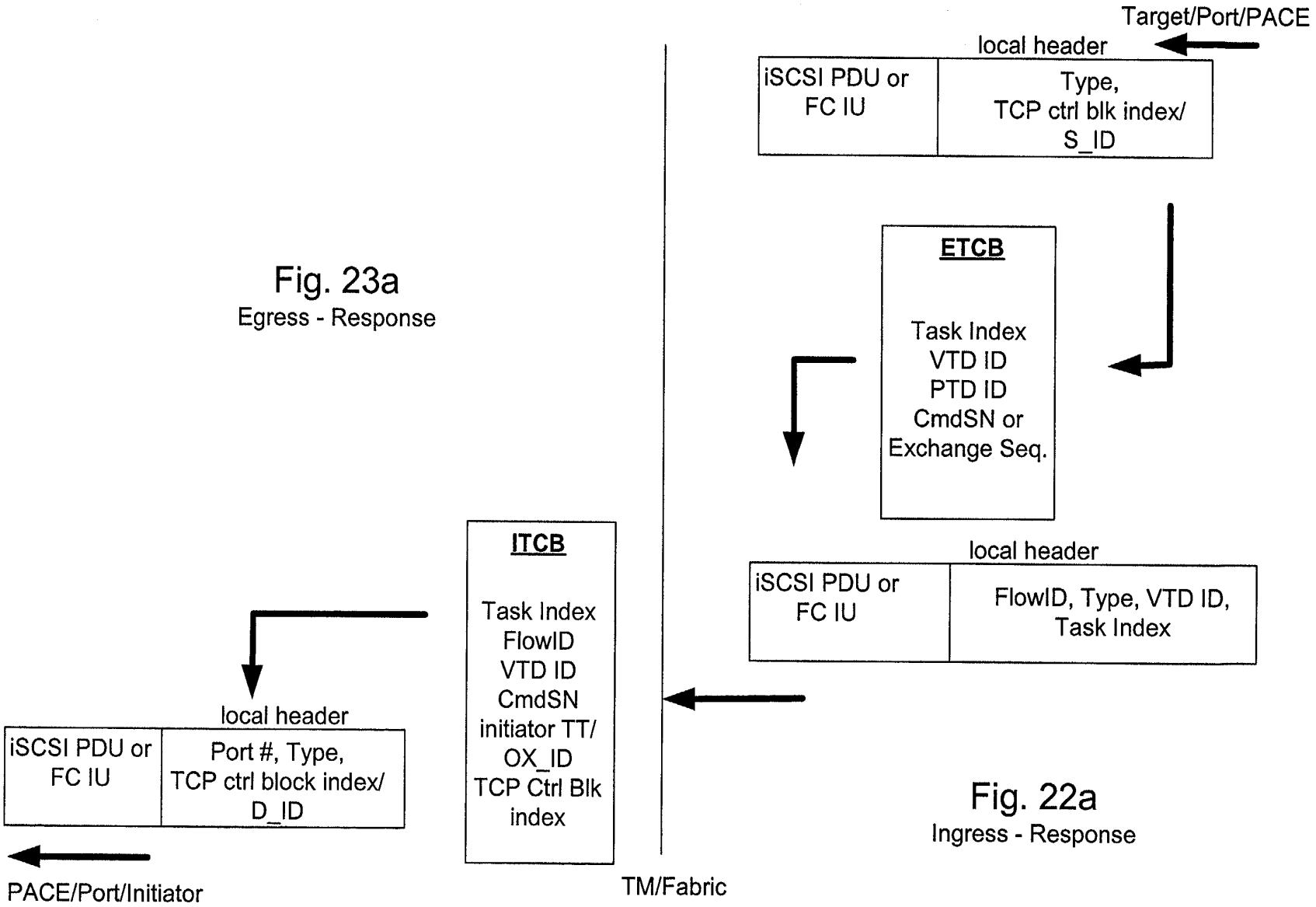


Fig. 23a
Egress - Response

Fig. 22a
Ingress - Response

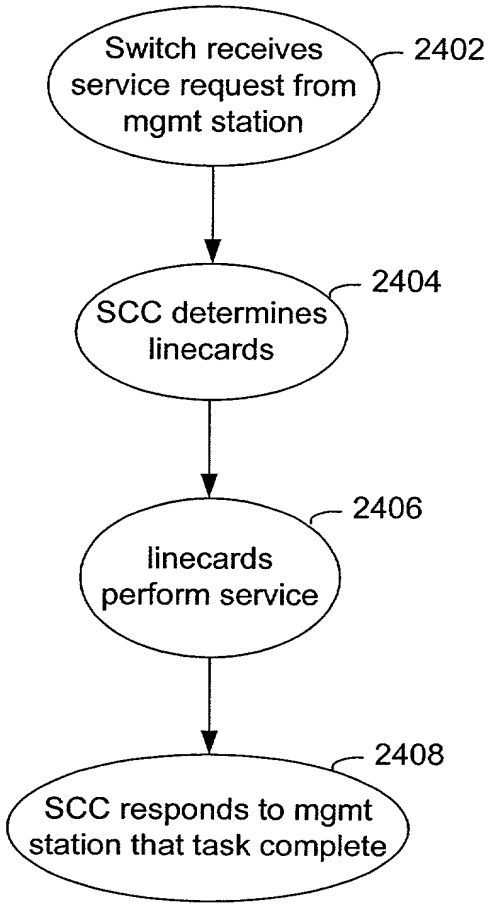

Switch receives service request from mgmt station
2402
SCC determines linecards
2404
linecards perform service
2406
SCC responds to mgmt station that task complete
2408

Fig.24

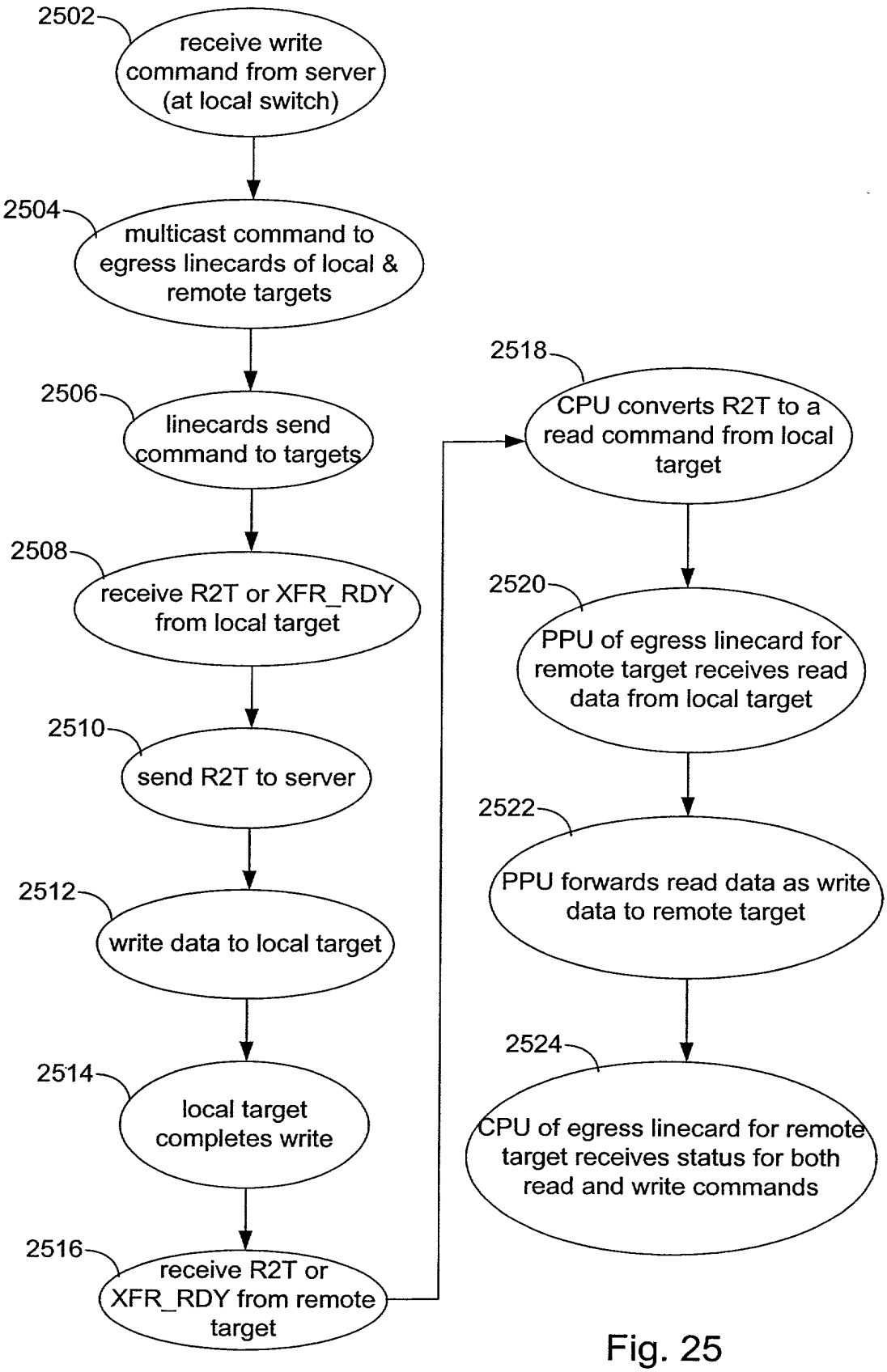

2502
receive write command from server (at local switch)
2504
multicast command to egress linecards of local & remote targets
2506
linecards send command to targets
2508
receive R2T or XFR_RDY from local target
2510
send R2T to server
2512
write data to local target
2514
local target completes write
2516
receive R2T or XFR_RDY from remote target
2518
CPU converts R2T to a read command from local target
2520
PPU of egress linecard for remote target receives read data from local target
2522
PPU forwards read data as write data to remote target
2524
CPU of egress linecard for remote target receives status for both read and write commands

Fig. 25

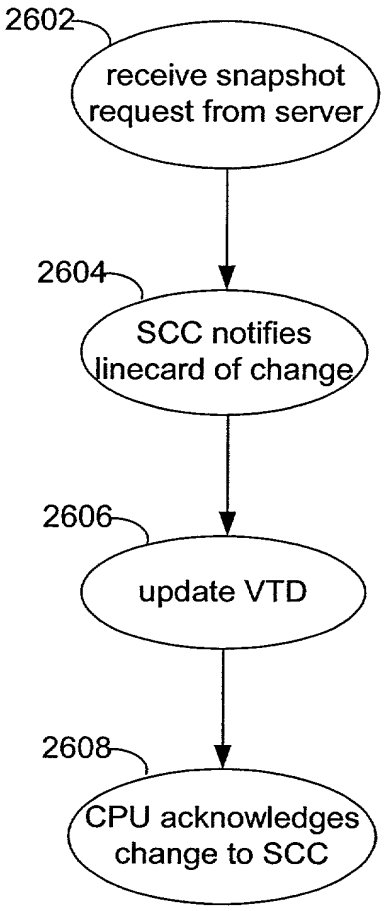
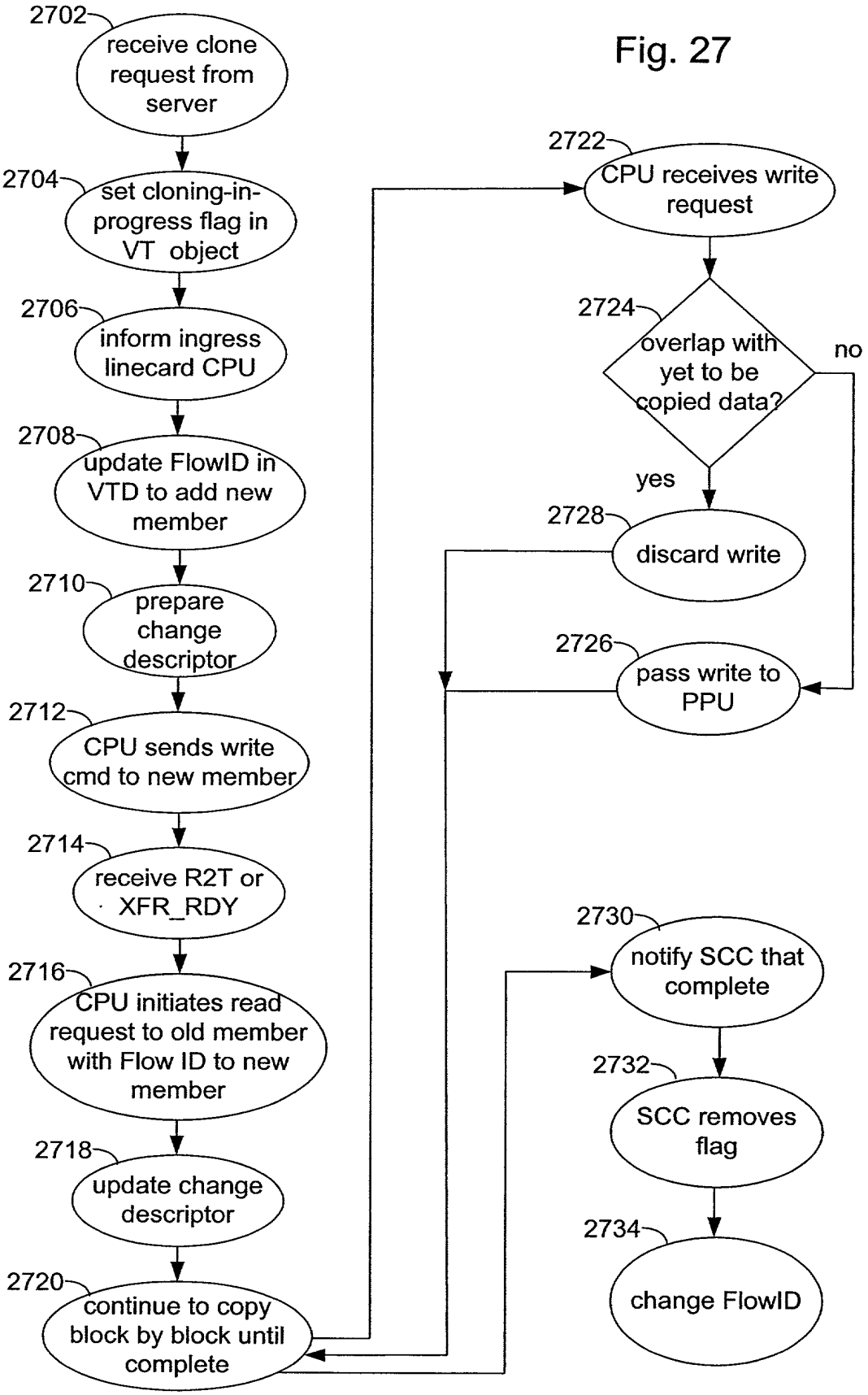

2602
receive snapshot request from server
2604
SCC notifies linecard of change
2606
update VTD
2608
CPU acknowledges change to SCC

Fig. 26

Fig. 27


2702
receive clone request from server
2704
set cloning-in-progress flag in VT object
2706
inform ingress linecard CPU
2708
update FlowID in VTD to add new member
2710
prepare change descriptor
2712
CPU sends write cmd to new member
2714
receive R2T or XFR_RDY
2716
CPU initiates read request to old member with Flow ID to new member
2718
update change descriptor
2720
continue to copy block by block until complete
2722
CPU receives write request
2724
overlap with yet to be copied data?
yes
no
2728
discard write
2726
pass write to PPU
2730
notify SCC that complete
2732
SCC removes flag
2734
change FlowID

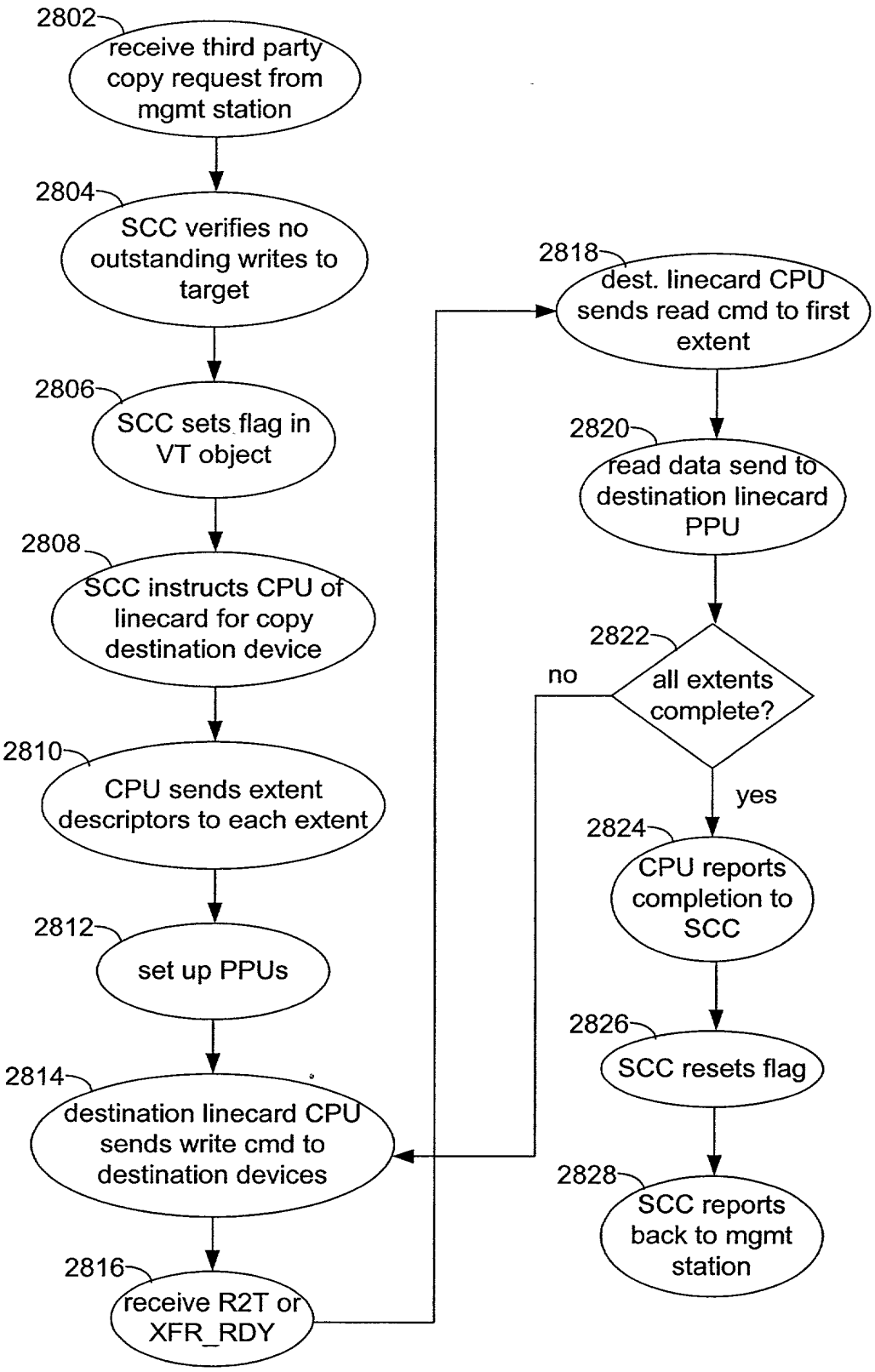

2802
receive third party copy request from mgmt station
2804
SCC verifies no outstanding writes to target
2806
SCC sets flag in VT object
2808
SCC instructs CPU of linecard for copy destination device
2810
CPU sends extent descriptors to each extent
2812
set up PPUs
2814
destination linecard CPU sends write cmd to destination devices
2816
receive R2T or XFR_RDY
2818
dest. linecard CPU sends read cmd to first extent
2820
read data send to destination linecard PPU
2822
all extents complete?
no
yes
2824
CPU reports completion to SCC
2826
SCC resets flag
2828
SCC reports back to mgmt station

Fig. 28